NTIS# PB98-110570

Final Report

AN ANALYSIS OF THE VALIDITY OF THE DISCRETIONARY COMPONENT OF DIAGNOSTIC COST GROUP ADJUSTERS

Frank W. Porell, Ph.D.

DataChron Health Systems, Inc. and
University of Massachusetts at Boston

Leonard Gruenberg, Ph.D.

DataChron Health Systems, Inc.

DataChron Health Systems, Inc.
Cambridge, MA 02139

July, 1997

Acknowledgments

This research was supported by a Cooperative Agreement No. 17-C-90295\01 from the Health Care Financing Administration. The views and opinions expressed in the report are the authors and do not necessarily represent those of HCFA. The authors would like to acknowledge the assistance of Alan Malbon in analytic file construction and Robert Boutwell for his guidance in issues concerning the clinical interpretations of diagnostic codes. Thanks are extended to Lev Shlimovich and Mary Whelan for their tedious work associated with mapping geographic units. We are grateful to Connie Tai for her help in the statistical analysis, and to Priyanthi Silva, Xiaoping Zhou, Terri Salmons, and Dianne Cleary for their editorial help and assistance in production of the report. Finally, we are grateful to Micheal Kendix for his patience and for his comments upon earlier an earlier draft of this report.

AN ANALYSIS OF THE VALIDITY OF THE DISCRETIONARY COMPONENT OF DIAGNOSTIC COST GROUP (DCG) RISK ADJUSTERS

EXECUTIVE SUMMARY

BACKGROUND

The objective of this study is to investigate the validity of the discretionary component of the DCG risk classification system of Ellis and Ash (1995). The discretionary ratings employed in the DCG model of Ellis and Ash (1995) potentially serve a very important function since increased capitation rates (over a base group level) are not made for prior hospitalizations with diagnoses that are in one manner or another deemed to be "discretionary." A primary reason for excluding such discretionary admissions in the prior use DCG risk adjusters is the belief that a significant part of the observed lower hospital utilization rates of HMOs may be attributed to HMOs being more effective than the FFS sector in avoiding such admissions. The failure to remove discretionary admissions associated with HMO-FFS hospitalization practice patterns would penalize HMOs that were successful in controlling discretionary hospital use. This study tests the validity of the DCG discretionary ratings by assessing:

- **the extent to which discretionary hospital admissions account for differences in the hospital admission rates of Medicare health maintenance organization (HMO) enrollees and Medicare fee-for-service beneficiaries;**
- **the extent to which discretionary hospital admissions account for geographic variations in the hospital admission rates of Medicare fee-for-service beneficiaries.**

To date, there has not been an empirical appraisal of the effectiveness of the discretionary component of the DCG risk classification model of Ellis and Ash (1995) in singling out hospitalizations more likely to be reduced by Medicare risk HMOs. Such an appraisal is important given the recent development of "hierarchical co-existing condition" (HCC) models by Ellis, et. al (1995) as a refinement to earlier DCG models. Whereas the earlier DCG models of Ash, et al. (1989) and Ellis and Ash (1995) address the problem of potential biases associated with HMO-FFS differences in medical practice styles by *limiting* prior hospitalizations to nondiscretionary ones, Ellis et. al (1995) strive to avoid such potential biases in HCC models by *expanding* the sources of prior use diagnostic information to

include all recorded diagnostic codes, including secondary hospital diagnoses, and diagnoses recorded on ambulatory outpatient and physician claims data.

While the question of whether use of all recorded patient diagnoses for risk classification is a better way to deal with such potential payment rate biases than the exclusion of highly discretionary hospitalizations has obvious importance, the pros and cons cannot be meaningfully evaluated without an empirical appraisal of the effectiveness of the DCG discretionary classifications in distinguishing "high discretion" hospitalizations. This study provides such an appraisal.

RESEARCH APPROACH

The empirical analysis of the validity of the DCG discretion classifications is founded on empirical tests of two main hypotheses. Under the general theoretical expectation that HMOs should be more successful than the FFS in reducing discretionary hospital admissions, it is formally hypothesized that:

- **"high discretion" hospital admissions will account for a significant part of the higher overall hospital admission rates of FFS Medicare beneficiaries relative to Medicare risk HMO enrollees in the same geographic markets.**

In addition to HMO-FFS comparisons, the diagnostic composition of Medicare FFS hospitalizations will be compared among geographic areas with higher and lower rates of Medicare hospital use. The most widely held view is that widely documented geographic variations in medical care use are largely the result of medical practice style differences (Wennberg et. al 1982; Wennberg 1985,1987; Wennberg et. al 1987, Wennberg et. al 1989). Employing logic similar to that raised above about HMO-FFS practice style differences, geographic differences in medical practice styles and the propensity to hospitalize should also be reflected in observable geographic differences in the diagnostic composition of inpatient hospitalizations. It is formally hypothesized that:

- **"higher discretion" hospital admissions will account for a significantly greater part of overall hospital admission rates in high-use geographic markets than in low-use geographic markets.**

The empirical performance of the DCG discretionary ratings are compared with the performance of several alternative discretion classifications found in the health services research literature. Anderson et. al (1989) developed two discretion classification systems derived from physician ratings of diagnoses in which the general concept of variation in the likelihood of a patient with certain medical conditions to be hospitalized was distinguished as resulting from variation in the severity level of the patient and physician discretion toward alternative treatment of the medical conditions. In order to provide some insight about the potential limits of using diagnostic information to classify hospitalizations into discretion classes, the empirical performance of the DCG and Anderson discretion classifications are compared to the index of discretion developed by Roos et. al (1988). In contrast to the a priori physician ratings of discretion in the DCG model of Ellis and Ash (1988) and those of Anderson et. al (1989) based on clinical judgement, the discretion index of Roos et. al (1988) was empirically derived based on consistent patterns in the geographic variability of hospital discharge rates among geographic in several states. Accordingly, the empirical performance of the Roos index of discretion may be viewed as a likely upper limit for discretion classifications based on diagnostic information.

PRINCIPAL FINDINGS

HMO-FFS Differences in Discretionary Hospitalizations

Descriptive analyses of the diagnostic composition of over 2.6 million Medicare HMO and FFS hospitalizations for four states provided little empirical support for the main study hypothesis, namely, that a significant portion of lower Medicare risk HMO hospital use rates are associated with their success in reducing discretionary hospitalizations as defined by the DCG discretion classes. Rather, the diagnostic composition of Medicare risk HMO and FFS hospitalizations appear to be relatively invariant with respect to both DCG risk class and discretion score classifications. As similar results were found using the alternative discretion classifications of Anderson, et al. (1989) and Roos, et al. (1988), the DCG model findings are unlikely to be the result of the misclassification of a few diagnoses among discretion classes.

Geographic Variations in Hospital Use and Discretionary Hospitalizations

Geographic analyses were conducted to test the main study hypotheses regarding the relative geographic variability of high versus low discretion hospitalizations, and the degree to which higher overall rates of hospital use are attributable to excessive high discretion hospitalizations. Correlation and regression analysis findings for the DCG and alternative discretion classifications suggest higher use rate geographic areas tend to have a greater share of hospitalizations classified as high discretion, but higher overall hospital use rates are not largely attributable to excessive discretionary hospital use.

Our empirical analyses suggest that higher overall hospital use rates among geographic areas are more strongly associated with a residual group of hospitalizations for all conditions not classified as "low discretion." Geographic areas with high or low overall hospital use rates appear to have relatively similar absolute hospital use rates for low discretion hospitalizations. A low fraction of "low discretion" hospitalizations in a geographic areas appears to serve as a much more effective marker for distinguishing high overall use rate geographic areas than a high fraction of "high discretion" hospitalizations.

Comparisons of the relative empirical performance of the DCG and Anderson, et al. (1989) discretion classifications in the correlation and regression analyses of Chapter 4 indicate only very modest differences among alternative discretion classifications derived from physician a priori ratings of principle diagnoses of hospitalizations. The empirically-derived Roos, et al. (1988) index of discretion stood apart from the DCG and Anderson discretion classifications in terms of superior empirical performance. Nevertheless, the empirical findings from the Roos index of discretion could hardly lead one to conclude that higher overall hospital use rates were predominantly the result of highly discretionary hospital use. The relative geographic invariance of the composition of hospitalizations under the Roos index of classification suggests that there are significant limitations associated with distinguishing the discretion level of hospitalizations from the principal discharge diagnostic information.

POLICY IMPLICATIONS FOR PRIOR USE MODEL RISK CLASSIFICATION

The study findings have direct implications toward the approach Ellis and Ash (1995) took for reducing the potential biases of medical practice style differences on prior use risk classifications. Our empirical analyses indicate that it is very difficult to distinguish well-defined subgroups of hospitalizations which account for significant portions of observed HMO-FFS and geographic differences in hospital use rates on the basis of diagnostic information from claims data. In general, the study findings suggest that the discretionary component of the DCG model of Ellis and Ash (1995) is unlikely to serve its original intended purpose. It appears that either systematic HMO-FFS practice style differences have little impact on diagnostic risk classifications of Medicare beneficiaries derived from inpatient hospitalization data, or the measurement of physician discretion in the DCG model of Ellis and Ash (1995) and other existing alternative discretion classification systems do not have sufficient validity to warrant the exclusion of certain "high discretion" hospitalizations for purposes of higher risk classification. Our findings provide empirical support for Ellis, et al.'s (1995) abandonment of the concept of physician discretion for excluding certain hospitalizations for assignment of enrollees to higher risk cells in their development of HCC models. However, they do not have any direct implications toward the merits of expanding sources of diagnostic information in HCC risk classification to ambulatory claims. Potential concerns over biases associated with HMO-FFS practice style differences or provider gaming behavior with respect to HCC models should be focused on diagnostic assignment derived from outpatient and/or physician utilization claims.

IMPLICATIONS FOR SMALL AREA ANALYSIS OF GEOGRAPHIC VARIATIONS

The study findings have some important broader implications for studies of geographic variations in hospital use as well. While concerns over the reliability of primary payer fields in the hospital discharge data precluded direct analyses of Medicare risk HMO and FFS differences in hospital utilization rates, our limited use of Medicare risk HMO enrollment data in the empirical geographic analyses of hospital use rates for combined Medicare risk HMO and FFS beneficiary populations suggests that estimation of disaggregated HMO and FFS hospital use rate models would be

likely to produce some valuable empirical insight about the impacts of population health status differences on variations in hospital utilization rates. Our descriptive analyses indicated that risk HMO-FFS differences in estimated age-sex adjusted mortality rates were much greater than geographic differences in mortality rates for the combined HMO and FFS study populations, particularly in California and New York. Given the strong correlations we found between hospital use rates and the mortality rates for the combined HMO and FFS populations of geographic areas, the greater dispersion in mortality rates among Medicare risk HMO versus FFS beneficiaries may provide a means for distinguishing population health status effects on geographic variations in hospital use rates that are muted by the aggregation of HMO and FFS data.

AN ANALYSIS OF THE VALIDITY OF THE DISCRETIONARY COMPONENT OF DIAGNOSTIC COST GROUP (DCG) RISK ADJUSTERS

TABLE OF CONTENTS

TABLE OF CONTENTS

(continued)

Appendix A METHODOLOGY FOR DEFINING GEOGRAPHIC UNITS

TESTING THE VALIDITY OF THE DISCRETIONARY COMPONENT OF DIAGNOSTIC COST GROUP ADJUSTERS

INTRODUCTION AND BACKGROUND

CHAPTER 1

1.1 INTRODUCTION

The Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982 established that capitation payments to health maintenance organizations (HMOs) entering into risk contracts with the Medicare Program were to be set on the basis of the expected fee-for-service (FFS) Medicare reimbursements of HMO enrollees. The expected FFS reimbursements of Medicare HMO enrollees have been determined operationally on the basis of the projected Medicare FFS reimbursement rates with the "adjusted average per capita cost" (AAPCC) formula (Kunkel and Powell 1981). Relative differences in the Medicare expenditure risks between Medicare HMO enrollees and non-enrollees residing in the same county are currently accounted for through a set of national underwriting risk factors based on age, gender, welfare status, and institutional status. Consistent findings of favorable selection bias in Medicare HMOs has drawn much attention to the inadequacy of these current AAPCC risk factor adjustments (Eggers and Prihoda 1982; Brown 1988; Porell and Turner 1990a,1990b; Riley et al. 1989; Brown et al. 1993; U.S. General Accounting Office 1986; Riley and Lubitz 1989).

Concerns over HMO selection bias spawned a considerable volume of research aimed at development of risk factors that would better reflect differences in the health status risks among Medicare beneficiaries. Most of this research has been focused on risk adjusters based on simple prior use measures such as the number of prior hospitalizations (e.g., Lubitz et al. 1985; Gruenberg et al. 1986; Epstein and Cumella 1988; Gruenberg et al. 1989; Newhouse et al. 1989), or more complex models incorporating diagnostic information about prior service use (Ash et al. 1989; Ellis and Ash

1995; Ellis et al. 1995). These studies have indicated significant improvements in predicting future Medicare reimbursements when prior use variables were added to demographic AAPCC risk factors.

A main drawback of simple prior use risk adjusters is that a significant portion of the prior use differences among individuals could reflect physician practice pattern differences as much as morbidity differences (Gruenberg et al. 1986; Ash et al. 1989). This limits their utility for updating enrollee risk classifications over time using actual HMO utilization. To the extent that HMOs are relatively successful at reducing certain kinds of hospitalizations by substitution of outpatient for inpatient care, a risk classification based on all prior hospitalizations would essentially penalize HMOs that are successful in controlling hospital utilization in this way. That is, an HMO would only receive a higher subsequent capitation rate for a Medicare enrollee treated for a medical condition if treatment was provided in an inpatient setting.

The limitations of simple prior use measures stimulated the development of the diagnostic cost group (DCG) models of Ash et al. (1989) and Ellis and Ash (1995). While DCG risk models are still prior use models, diagnostic codes are used in two important ways which address some shortcomings of simpler prior use models. First, diagnostic codes are used to differentiate among prior hospitalizations with differing levels of expected future costs. Enrollees who were hospitalized for a condition with higher levels of expected future costs would be assigned to risk cells with higher capitation rates. Second, hospitalizations with principal diagnoses for which physicians exercise considerable discretion in decisions to hospitalize are not used for assignment of individuals to higher paying risk cells. This treatment of "highly discretionary" hospitalizations in the DCG model serves an important purpose since it should insulate inpatient versus outpatient treatment decisions from the influence of capitation reimbursement incentives.

To date there has not been an empirical appraisal of the effectiveness of the discretionary ratings of the DCG risk classification model in singling out hospitalizations which are more likely to be

reduced by Medicare HMOs relative to FFS providers because of differences in medical practice style. Such an appraisal is important given recent DCG model refinements known as "hierarchical co-existing condition" (HCC) models (Ellis, et al. 1995). Whereas potential biases associated with HMO-FFS practice style differences are addressed in the DCG model by *limiting* prior use risk classification to nondiscretionary hospitalizations, Ellis et al. (1995) seek to avoid such biases by *expanding* the sources of diagnostic codes for risk classification to include secondary hospital diagnoses, and diagnoses recorded on ambulatory outpatient and physician claims data. In either case a Medicare HMO enrollee's risk classification should not vary because of discretion in an HMO's choice of inpatient or outpatient treatment for certain medical conditions.

Whether potential biases in risk classification due to medical practice style variations are better handled by using all recorded patient diagnoses from all claims, or by limiting diagnoses to the principal diagnoses of nondiscretionary hospitalizations is an important question. A practical answer will depend very much upon the validity of the DCG discretionary ratings used to distinguish "highly discretionary" hospitalizations. If the DCG discretionary ratings of Ellis and Ash (1995) are not valid in the sense that they do not effectively distinguish hospitalizations for which there is a high level physician discretion toward inpatient hospitalization, it is difficult to defend differential treatment of the identified "highly discretionary" hospitalizations for purposes of risk classification.

1.2 STUDY GOALS AND HYPOTHESES

The major goal of the study is to investigate the validity of the discretionary component of the DCG risk classification system through a comprehensive empirical analysis of differences in the diagnostic composition of hospitalizations of Medicare beneficiaries who receive care in settings where medical practice styles are likely to differ. Unless the level of physician discretion toward inpatient versus alternative treatment setting doesn't vary much among diagnosed medical conditions, or the

measurement of discretion toward hospitalization is invalid, or HMO-FFS health status differences have not been adequately controlled for, systematic HMO-FFS differences in the propensity to hospitalize for certain medical conditions ought to be reflected in observable HMO-FFS differences in the diagnostic composition of inpatient hospitalizations.

Under the general theoretical expectation that HMOs should be more successful than the FFS in reducing discretionary hospital admissions, it is formally hypothesized that:

- **"high discretion" hospital admissions will account for a significantly greater portion of total hospital admissions for FFS Medicare beneficiaries relative than for Medicare risk HMO enrollees in the same geographic markets.**

In addition to HMO-FFS comparisons, the diagnostic composition of Medicare hospitalizations will be compared among geographic areas with higher and lower rates of Medicare hospital use. The most widely held view is that widely documented geographic variations in medical care use are largely the result of medical practice style differences (Wennberg et al. 1982; Wennberg 1985,1987; Wennberg et al. 1987, Wennberg et al. 1989). Employing logic similar to that raised above about HMO-FFS practice style differences, geographic differences in medical practice styles and the propensity to hospitalize should also be reflected in observable geographic differences in the diagnostic composition of inpatient hospitalizations. It is formally hypothesized that:

- **"higher discretion" hospital admissions will account for a significantly greater portion of total hospital admission in high-use rate geographic markets than in low-use rate geographic markets.**

Since the original intent of the discretionary component of the DCG model of Ellis and Ash (1995) was to ensure risk classifications that were relatively invulnerable to medical practice style differences generally, geographic comparisons of diagnostic composition can provide an important source of information regarding the validity of the DCG discretionary classifications with respect to their intended purpose.

It is also important to broadly assess the general limitations of discretion rating systems based solely on diagnostic information. For example, if very little, or any, of the higher overall rate of Medicare FFS hospital admissions relative to Medicare HMO hospital admissions can be attributed to greater rates of highly discretionary hospital admissions, this could reflect a limitation of the specific DCG discretionary ratings developed in Ellis and Ash (1995), or a more fundamental limitation in our ability to rate the discretion level of hospitalizations based solely on diagnostic information. To provide some insight on such issues, the relative empirical performance of the DCG discretionary ratings will be compared with the performance of the alternative physician discretion classifications of Anderson, et al. (1989), and the index of discretion of Roos, et al. (1988). Since the index of discretion of Roos, et al. (1988) was derived empirically from consistent observed (higher/lower) geographic variations in diagnostic-specific hospital admission rates among several states, it is expected to exhibit the best empirical performance with respect to the association between diagnostic composition of hospitalizations and variations in overall hospital use rates. Accordingly, its empirical performance should provide a basis for assessing the broader limitations of simple discretion classifications based solely on diagnostic information.

1.3 AN OVERVIEW OF CHAPTERS OF THE REPORT

The entire report contains six chapters including this one. The following chapter (Chapter 2) contains a descriptive analysis of the discretionary composition of Medicare risk HMO and FFS hospitalizations in the study states. It begins with a detailed discussion of the DCG risk classification model of Ellis and Ash (1995) and the development of its discretion classifications. This is followed by a detailed discussion of the study data sources and descriptive information about the demographic attributes, hospital use rates, and mortality rates of the study populations. The chapter concludes with a

descriptive analysis comparing the discretionary composition of risk HMO and FFS hospitalizations under the DCG discretion classification system.

The third chapter of the report contains a description of the alternative discretion classification systems of Anderson et al. (1989) and Roos, et al. (1988) and an analysis of the agreement between the DCG discretion classifications those of these alternative classification systems. In this chapter HMO-FFS differences in discretionary composition of hospitalizations are compared for the three alternative discretion classifications with a descriptive analysis similar to that of Chapter 2.

The fourth chapter of the report is devoted to analyses of geographic variations in overall hospital use rates and their association with the discretionary composition of hospitalizations. The chapter contains an overview of relevant geographic variations literature as well as a discussion of the methodology employed in the delineation of study geographic units and descriptive statistics about the study populations. The chapter concludes with a series of geographic analyses relating overall hospital use to discretionary composition of hospitalizations for each of the four alternative discretion classifications.

The fifth chapter contains empirical findings from a multivariate analysis which extends the analyses reported in Chapter 4. The classifications of high and low discretion hospital admissions are further tested through a simple econometric model in which health delivery market supply variables are related to hospital use rates for high discretion versus low discretion hospital discharges in a multivariate regression models.

The last chapter of the report (Chapter 6) contains a summary of the study findings, a discussion of their policy implications toward risk classification model refinements, and potential directions for future research.

A COMPARISON OF DCG DISCRETION LEVELS IN MEDICARE RISK HMO AND FEE-FOR-SERVICE HOSPITALIZATIONS

CHAPTER 2

2.1 INTRODUCTION

There is fairly ample empirical evidence that hospital utilization rates of Medicare HMO enrollees are lower than those of FFS nonenrollees. While much of the difference in utilization has been attributed to the selection bias in HMO enrollment rather than control of utilization, it is reasonable to posit that HMOs may be more effective in reducing hospitalizations for which there is greater discretion toward inpatient versus outpatient treatment. If such discretionary hospitalizations can be identified through principal diagnosis information, HMO-FFS differences in utilization control ought to be reflected in HMO-FFS differences in the diagnostic composition of hospitalizations.

Currently there is relatively little knowledge about the diagnostic composition of Medicare HMO hospital use other than that provided by research conducted by some HMO researchers (e.g., Hornbrook et al. 1988) and data from selected HMO demonstration sites. The only published study to date which has actually investigated compositional differences in HMO and FFS hospitalizations was conducted by Luft (1978), who pieced together data from multiple sources on HMO and FFS overall hospital admission rates and rates for eight specific procedures or diagnoses suggested in the literature as being relatively discretionary (e.g., hysterectomy, tonsillectomy, etc.). While some attempts were made to account for HMO-FFS health status differences through age-sex adjustments, the data did not support the expectation that HMOs would derive a disproportionate share of their reductions in total hospital admissions from more discretionary admissions. Rather the data strongly showed that HMOs had lower admission rates across the board for all kinds of hospitalizations.

Given the evolution and expansion of managed health care over the last 20 years, the medical practice style of physicians in the prepaid group practices studied by Luft (1978) may be quite different from those in contemporary HMOs. It is unclear whether Luft (1978) would find similar unexpected results today. The purpose of this chapter is to investigate HMO-FFS differences in medical practice styles through a comparative analysis of the diagnostic composition of Medicare HMO and FFS hospitalizations. More specifically, we will compare the discretion ratings and the associated DCG risk classifications (as developed in Ellis and Ash (1995)) of Medicare HMO and FFS hospitalizations for the four study states. Given the financial incentives of capitation, disproportionately fewer highly discretionary hospitalizations are expected for Medicare risk HMO enrollees.

Since HMOs with Medicare risk contracts are not required to submit hospital claims that can be directly compared to Medicare FFS hospital claims, there is a paucity of empirical evidence on the composition of Medicare HMO hospitalizations. In this study primary payor information from state hospital discharge data files from four states is used to distinguish Medicare FFS and HMO hospitalizations for patients 65 years old or older. The principal diagnosis recorded on the discharge records are used for assigning hospitalizations to DCG and discretion score categories. The descriptive analyses reported in this chapter permit an assessment of the effectiveness of the physician discretion component of the DCG model of Ellis and Ash (1995) in singling out the kinds of hospitalizations Medicare HMOs have been successful in avoiding relative to the FFS providers.

2.2 BACKGROUND

2.2.1 The DCG Risk Classification Model

Ash, et al. (1989) and Ellis and Ash (1995) employed diagnostic information from prior hospitalizations to develop a set of health status-based risk classifications known as DCG models. The DCG risk classifications, based on the principal diagnosis of inpatient hospitalizations, are intended to

reflect health status differentials, in the sense that persons at greater risk of high future medical care expenses are distinguished as having poorer health status than persons at lesser risk of high future medical care expenses. For example, Medicare beneficiaries hospitalized in the last year with a primary diagnosis of a cancer such as malignant neoplasm of the liver (ICD-9 code 155) have much higher than average expected annual Medicare reimbursements. On the other hand, beneficiaries who were not hospitalized at all in the last year, or who were hospitalized with a primary diagnosis such as genital prolapse (ICD-9 code 618), have lower than average expected annual Medicare reimbursements.

Ellis and Ash (1995) distinguished seven ordinal DCG risk classes (DCGs 1-7) with higher expected future Medicare reimbursements than a base category DCG 0. The DCG 0 groups was comprised of persons who were not hospitalized in the previous year (the largest group), persons whose hospitalizations were not found to be associated with significantly higher future Medicare reimbursements, persons with hospital stays shorter than 3 days, and persons who were hospitalized with diagnoses judged by physicians to be highly discretionary. Reimbursement ratios, defined as the ratio of expected subsequent year Medicare costs for a subgroup of persons to the population average, ranged from 0.89 for the base class DCG 0 to 5.96 in the highest risk class DCG 7.

Since decisions to hospitalize may often be subject to the influences of different medical practice styles and beliefs, a conscious decision was made by Ellis and Ash (1995) to exclude "highly discretionary" hospitalizations for risk classification purposes in the empirical development of the DCG model. Highly discretionary hospitalizations were identified by asking a panel of three physicians to rate individual three or four digit ICD-9 diagnostic codes on the basis of clinical judgement concerning various forms of discretion. First, physicians were asked to rate diagnoses as to whether there was likely to be high, moderate, or low discretion on the part of physicians as to hospitalize a patient versus outpatient treatment, or possibly no treatment. Second, the panel was asked to designate diagnoses as

having "overstay potential" if there was a reasonable likelihood that an HMO could admit a patient inappropriately, and/or prolong a length of stay to three days to increase subsequent capitation revenues. Finally, physicians were asked to designate as "ambiguous" those diagnoses which were most vulnerable to coding practices, or where a vague or unspecific code could be used instead of a more specific diagnosis. The physicians were not supplied with any empirical information (e.g., readmission rates) to assist them in rating the diagnostic codes other than information about the relative frequency of hospitalizations by principal diagnosis. More specifically, physicians were given a list of principal diagnoses for Medicare hospitalizations with an national volume of 1,000 admissions or more per year.

The individual physician ratings of diagnoses for each of the three basic dimensions of discretion were combined together in the form of a simple numerical discretion rating through the following formula:

- 1 point was assigned to diagnoses with a low substitutability rating (i.e., high discretionary rating), and no other problem was designated;
- 2 points were assigned to diagnoses with a moderate substitutability rating, and no other problem was designated;
- 3 points were assigned to diagnoses with only a high substitutability rating, or a designation of overstay potential or ambiguous coding;
- 4 points were assigned if more than one of the problems of high substitutability, overstay potential, or ambiguous coding was noted.

A total discretion score ranging from 3 to 12 for each 3 or 4-digit ICD-9 diagnosis was then computed as the sum of the three physician scores. The minimum score of 3 would result if all physicians rated the diagnosis as have low substitutability potential and no other problem was designated. The maximum score of 12 would result when each physician noted two or more discretion problems. Table 2.1 contains a lists of selected 3-digit primary diagnoses among the 51 most frequent principal diagnoses for Medicare FFS hospitalizations in 1985 classified into the three discretion level categories

Table2. 1: DCG Discretion Score Ratings for Selected Common 3-digit ICD-9 Codes

Discretion Category	3-digit ICD-9 code
Low Discretion (3-4)	
Acute Myocardial Infarction	410
Malignant Neoplasm colon	153
Intestinal Obstruction	560
Pneumonia, Organism Nos	486
Malignant Neoplasm, Trachea/lung	162
Moderate Discretion (5-8)	
Asthma	493
Hyperplasia of Prostate	600
Diverticula of Intestine	562
Inguinal hernia	550
Thrombophlebitis	451
High Discretion (9-12)	
Heart Failure	428
Chronic Airway Obstruct Nec	496
Essential Hypertension	401
Gastrointestinal Hemorrhage	578
Cataract	366

advocated by Ellis and Ash (1995). "High discretion" hospitalizations were defined a by total discretion score of 9 or more. "Moderate discretion" and "low discretion" hospitalizations were defined by total discretion scores ranging between 5-8 and 3-4, respectively.

In sum, the DCG risk classification model of Ellis and Ash (1995) provides two useful ordinal scales for comparing the diagnostic composition of Medicare HMO and FFS hospitalizations. HMO-FFS comparisons of the distribution of hospitalizations among the eight ordinal DCG risk classifications (DCG0-DCG7) provides a simple, useful yardstick for drawing inferences about the health status differences in Medicare HMO and FFS populations. HMO-FFS comparisons of the distribution of hospitalizations by discretion classes provides a basis for assessing the effectiveness of these DCG discretion ratings in singling out hospitalizations that were disproportionately reduced by HMO providers relative to FFS providers.

2.3 DATA AND METHODOLOGY

2.3.1 Data Sources

The main source of data for the analysis were public use hospital discharge files for the 1992 calendar year for the states of California, Florida, Massachusetts, and New York. Collectively, these four states accounted for about 60 percent of Medicare risk HMO enrollment in July, 1992. The main factor governing the selection of these study states was the availability of public use hospital discharge data files with fields for patient age, patient gender, and primary payer. Age and primary payer source were crucial data elements for distinguishing hospitalizations that were very likely those of Medicare risk HMO enrollees. The other main factors governing the selection of study states included: level of Medicare risk HMO market penetration, and the level of geographic variation in hospital use rates. Two of the selected study states (California, and Florida) had relatively high Medicare risk HMO market penetration in 1992, while the others had relatively low market penetration rates at the state

level (Massachusetts, and New York). Regarding regional variations in hospital use rates, the states of Massachusetts and New York have reputations as states with high hospital use rates and longer stays, while California is notable for its lower hospital use rates. Basic descriptive statistics about the aged Medicare beneficiary populations in the four states will be presented in the empirical results section.

Given the paucity of available HMO hospital use data, public discharge files provide a possibly overlooked opportunity to gain some needed insight about hospitalizations of Medicare risk HMO enrollees. Since HMOs serve as the primary payer for Medicare risk contract enrollees and fiscal intermediaries do not reimburse hospitals for Medicare risk contract enrollees, the great bulk of hospitalizations of patients 65 years or older with the primary payer recorded as HMO should be hospitalized Medicare risk HMO enrollees. Little is known about the reliability of the specific discharge record fields that are used for determining Medicare eligibility and HMO enrollment status of hospitalized patients and there are some limitations of these data that should be noted.

First, hospitalized patients who are entitled to Part A of Medicare under Old Age Survivor Insurance (OASI) can only be approximated by selecting hospitalizations where patients are 65 years old or more. There is a relatively small but nonzero fraction of aged population which is not entitled to Part A of Medicare (i.e., mostly federal retirees) that will be included by selecting hospitalizations of Medicare beneficiaries based on the age of the patient. Second, and more importantly, Medicare HMO risk enrollment status of aged hospitalized patients can only be determined by the expected primary payer reported by hospitals, and in two of the study states (Massachusetts and New York) secondary payer. Unfortunately, little knowledge that exists about the reliability of expected payer source data reported by hospitals, and study resources did not permit a reliability analysis of the payer fields for the study states.

A third limitation of these data is that hospital discharges for individual HMOs cannot be distinguished from one another, nor can discharges of aged patients enrolled in HMOs with Medicare

cost contracts, including Health Care Prepayment Plan (HCPP) Part B-only cost contracts, be distinguished from discharges of Medicare FFS beneficiaries. Medicare holds primary payer status for cost HMO enrollees and their Part A hospital inpatient claims are processed along with those of Medicare FFS beneficiaries by fiscal intermediaries. Unless the practice styles of Medicare cost HMOs are more similar to those of Medicare FFS providers than Medicare risk HMOs, the mix of both cost HMO enrollees and FFS beneficiaries may diffuse some HMO-FFS differentials associated with incentives of capitated payments. Fortunately for the purpose of this study, Medicare cost HMO enrollment is relatively low in comparison to both Medicare risk HMO and Medicare FFS beneficiary populations in the study states.

Finally, there is a border crossing limitation of state discharge databases. These files do not include admissions of Medicare state residents to out-of-state hospitals. Recent research by Yip and Luft (1993), which compared Medicare discharge data with state hospital discharge data, found that border crossing was not a serious problem for two of the study states (California and New York). Only 4.4 percent of New York and 2.1 percent of California patients received care out-of-state.

Given these limitations of state discharge databases, some caution is warranted in interpretation of the empirical findings. However given the large HMO-FFS differences age-sex adjusted hospital admission rates, state discharge data should have sufficient reliability for assessing whether differences in the diagnostic composition of Medicare HMO and FFS hospitalizations are large enough such that discretionary hospital use can account for much of the higher Medicare FFS hospital utilization rates.

The other major data source for this descriptive analysis of HMO-FFS hospital use is the Denominator file from the Medicare Statistical System. It was used to obtain person-year counts beneficiary Part A eligibility, HMO enrollment status, and mortality data for the computation of age-sex adjusted hospital admission and mortality rates in the study states by HMO and FFS groups.

The main drawback of the denominator file data is that the beneficiary geographic identifiers (i.e., state, county, and zip code) pertain to March of year following the calendar year of eligibility information. For the 1992 Denominator file, the geographic information pertains to March, 1993. Since the elderly exhibit much lower geographic mobility rates than the non-elderly, this should not pose a major problem. Comparisons of 1991 and 1992 aged beneficiary population counts by county among the study states showed only very minor population changes. Since the 1991 Denominator file would not contain the geographic location of beneficiaries who turned 65 years old in 1992, the 1992 Denominator file was employed in the study.

2.3.2 Assignment of Hospitalizations to HMO versus FFS Sectors

As already noted, primary payor data fields noting the expected source of payment at the time of discharge were used to assign hospital discharges of aged patients to HMO versus FFS sectors. In all study states the simple procedure of mapping HMO and Medicare primary payor codes to Medicare risk HMO and Medicare FFS sectors, respectively, accounted for between 92 and 96 percent of total hospital discharges for aged patients. In the case of the states of Massachusetts and New York, the secondary payor code was used next for assignment if that secondary payor code was either Medicare or HMO.

A residual group of discharges which had neither Medicare or HMO as primary or secondary payor, accounting for between 4 to 8 percent of total aged discharges among states, had payor codes that were widely scattered among categories such as Medicaid, Blue Cross, workers compensation, self-pay, and other government. A simple methodology based on relative likelihood was employed to assign discharges from these residual payors to either the HMO or FFS sector. Data from the denominator files revealed that Medicare risk HMO enrollments were not uniformly distributed among zip codes within each state and that there were many 5-digit zip codes where no Medicare risk HMO

enrollees resided at all. It would not make sense to assign any unassigned hospital discharges in these zip codes to the HMO sector. This basic principal guided the assignment of residual other payor group hospital discharges to HMO versus FFS sectors as described below.

Assigning residual group hospital discharges to HMO versus FFS sectors on the basis of relative HMO and FFS beneficiary population size would entail an implicit assumption of equal hospital admission rates among Medicare risk HMO and FFS beneficiaries that is inconsistent with all available data. Hence assignments were made on the basis of the distribution of hospital discharges that were already assigned to either the HMO or FFS by payor code.

The assignment methodology entailed several steps. First, all discharges already assigned to either HMO or FFS sectors were stratified into subgroups by 5-year age class, gender, and 5-digit zip code of residence and the proportion of HMO discharges in each age-sex-zip code class was calculated to serve as an estimate of the likelihood that an unassigned patient of each age-sex-zip code class was a risk HMO enrollee. Second, each unassigned hospital discharge was randomly assigned a number between 0 and 1. If the random number was smaller than estimated HMO proportion for the relevant age-sex-zip code class, the unassigned discharged was classified as an HMO discharge; otherwise it was assigned to the FFS sector.[1]

2.3.3 Aggregation of Total DCG Discretion Scores

In terms of HMO-FFS practice style differences, the most important dimension of the three aspects of the DCG model discretion score is the substitutability of outpatient versus inpatient care. Unfortunately, we could only obtain the total summary discretion scores rather than individual

[1] Given the highly skewed spatial distributions of both Medicare risk HMO enrollment and assigned hospital discharges, there was very little difference in the assignments when beneficiary residence populations were used to assign residual unassigned discharges to HMO versus FFS sectors.

physician ratings about inpatient/outpatient substitutability for all diagnoses. Ellis and Ash (1995) reported individual physician scores for all three dimensions of discretion for the 51 most frequent 3-digit ICD-9 principal hospital diagnoses in 1985, accounting for over 62 percent of Medicare hospital admissions with a length of stay of three days or more. Analyses of these data indicate that an aggregated total discretion score scale of three categories: high discretion (9-12), moderate discretion (5-8), and low discretion (3-4), should provide a good approximation to a three-category inpatient/outpatient substitutability scale. A three-category scale of high, moderate, and low discretion categories was created from individual physician ratings of substitutability. High discretion was defined to exist when all three physicians rated a diagnosis as having high substitutability, or when at least two of them rated the diagnosis as high substitutability and the third physician rated it as moderate substitutability. Low discretion was similarly defined to exist when at least two physicians rated a diagnosis as low substitutability and the third rated it at least moderate substitutability. All other combinations of physician ratings were define to be moderate discretion. Scores from this inpatient/outpatient substitutability scale were compared to a three-unit scale based upon total discretion scores with three categories: high (a score of 9 or more), moderate (a score ranging between 5 and 8), and low (a score of 4 or less) discretion categories. For the 51 most frequent primary diagnoses there were only three cases where the scales differed. Accordingly, the three-category discretion scale was employed in the empirical analyses.

2.3.4 Aggregation of Ellis and Ash DCG Risk Classes

Ellis and Ash (1995) distinguished seven ordinal DCG risk classes (DCGs 1-7) with higher expected future Medicare reimbursements than a base category DCG 0, comprised mostly of individuals who were not hospitalized. Our assignments of hospitalizations to DCG risk classes were made solely on the basis of diagnostic codes without regard to length of stay. This was consistent with

the recommendations of Ash, et al. (1989) regarding the inclusion of hospital stays of two days or less and all subsequent DCG model refinement studies. In the presentation of our empirical findings, the 8 original DCG risk classifications of Ellis and Ash (1995) were collapsed into four aggregated DCG risk categories. The specific aggregations were defined in the following manner. The base level DCG 0 risk cell was comprised only of "high discretion" hospitalizations, hospitalizations with the lowest expected future Medicare costs, and no hospitalizations. The remaining three aggregated DCG risk classes were derived by grouping together the original DCG risk classes having similar reimbursement ratios, reflecting similar levels of expected future Medicare reimbursements.[2] The DCG risk classes 1 and 2 with reimbursement ratios of 1.99 and 2.21, respectively, were grouped together. DCGs 3 and 4 with reimbursement ratios of 2.41 and 2.79, respectively, were grouped together. Finally, DCGs 5, 6, and 7 with reimbursement ratios of 3.35, 4.39, and 5.96, respectively, formed a group of diagnoses with the highest expected subsequent year Medicare costs.

2.3.5 ICD-9 CM Coding Changes and Assignments to DCG and Discretion Score Classes

Since Ellis and Ash (1995) employed 1984 Medicare data for assignment of hospitalizations to DCG and discretion score class and the study data were from 1992, it was necessary to examine whether changes in the interpretation of ICD-9 codes over the intervening eight years would result in significant changes in the intended assignments of hospitalizations. Frequency distributions of the 5-digit ICD-9 codes for all principal diagnoses of hospitalizations in the combined pool of data for the four study states were compiled and sorted by their relative frequency. A physician clinical consultant compared the more frequent principal diagnoses with a cumulative record of ICD-9 code interpretation

---

[2] Reimbursement ratios are defined as the ratio of expected subsequent year Medicare costs for a subgroup of persons to the population average for a risk class. A reimbursement ratio of 1.5 for a population subgroup means that the expected reimbursements of this subgroup are 1.5 times higher than the mean level of expected reimbursements per capita for the entire population.

changes occurring between 1986 and 1992 compiled by the Health Care Financing Administration. While some coding changes were noted for certain diagnostic codes, it was concluded that changes in interpretation of ICD-9 codes would have negligible impact on DCG and discretion score classifications.

2.3.6 Research Approach

The main research questions to be addressed through a descriptive empirical analysis of study data pertain to HMO-FFS differences in beneficiary demographic characteristics, hospital utilization rates, mortality rates, and the diagnostic composition of hospitalizations. The first set of analyses pertain to differences between the study Medicare risk HMO and FFS populations in terms of their age-sex demographic composition, and their age-sex adjusted hospital utilization and mortality rates within the study states. This is followed by a comparison of Medicare HMO and FFS distributions of hospital discharges and hospital days among the DCG risk classes of Ellis and Ash (1995). This comparison should show the extent to which expected reductions in Medicare HMO hospital use are disproportionately represented in higher or lower DCG risk classes. Finally, we will assess the extent to which HMO-FFS differences in hospital utilization rates can be associated with higher/lower proportions of high discretionary hospital admissions as defined in the discretionary component of the DCG model of Ellis and Ash (1995). Under the general theoretical expectation that HMOs should be more successful than the FFS in reducing discretionary hospital admissions it is hypothesized that "high discretion" hospital admissions will account for a significant part of the expected higher overall hospital use rates of FFS Medicare beneficiaries relative to Medicare HMO enrollees.

2.4 EMPIRICAL RESULTS

2.4.1 The Study Population

Table 2.2 contains basic descriptive statistics about the size and composition of the aged Medicare populations in the four study states. Beneficiary populations are measured in terms of person-years of Medicare Part A eligibility. A casual examination of the table reveals some of the differences in demographic composition and risk HMO market penetration rates that guided the selection of these study states.

The age and gender composition of the aged Medicare populations reported in the top of Table 2.2 show the study states of California and Florida to differ from the Northeastern states of New York and Massachusetts. Males account for a higher percentage of the aged Medicare populations of California (41.7%) and Florida (42.1%) than for New York (39.1%) and Massachusetts (38.4%). While the differences by age class are fairly modest, the age distribution data indicate somewhat older beneficiary populations in the states of New York and Massachusetts relative to California and Florida. The most notable difference was in the oldest age category of beneficiaries aged 85 years or more, where the percentage of "oldest" old beneficiaries in New York (11.0%) and Massachusetts (11.2%) were substantially larger than that in California (9.6%) and Florida (9.2%). These demographic differences most likely reflect the effects of post-retirement outmigration patterns from the two Northeastern study states.

The HCFA Denominator files contains information about aged beneficiaries who were previously entitled to Medicare due to disability before turning 65 years of age. The data suggest there are relatively modest differences among the study states in the prevalence of such aged beneficiaries. On the other hand, the percent of aged beneficiaries with Medicaid buy-in status varied more among the states. It was particularly high in the state of California, where dually-eligible beneficiaries accounted for 13.6 percent of aged Part A person-years.

Table 2.2: Descriptive Statistics of the Study Population

	State of Residence				
Beneficiary Attributes	California	Florida	New York	Massachusetts	Total
Part A person-years (in thousands)	2960	2196	2235	806	8198
Gender					
% Males	11.8	42.1	39.1	31.6	40.8
% Females	58.3	57.9	60.9	61.6	59.2
Age Class					
% 65-69 yrs	34.0	32.4	32.7	31.6	33.0
% 70-74 yrs	25.9	26.0	24.8	25.3	25.6
% 75-79 yrs	18.7	19.7	18.8	19.1	19.0
% 80-84 yrs	11.8	12.7	12.7	12.8	12.4
& 85+ yrs	9.6	9.2	18.8	11.2	10.1
Percent Medicaid buy-in status	13.6	7.5	5.9	7.9	9.3
Percent with disability as original reason for entitlement	6.7	5.9	6.6	5.4	6.3
Percent Enrolled in Medicare Risk HMO	16.9	11.9	3.4	5.5	10.8
Percent enrolled in Cost HMO	8.0	0.5	2.7	1.7	3.9
Percent in fee-for-service	75.1	87.5	93.9	92.8	85.3

As noted earlier, Medicare risk HMO market penetration was substantially higher in California (16.9%) and Florida (11.9%) relative to New York (3.4%) and Massachusetts (5.5%). Cost HMO market penetration was relatively modest in all of the states except California where about 8 percent of aged Part A person-years were accounted for by cost HMO enrollment. The bulk of this cost HMO enrollment was concentrated in the Northern part of the state. To the extent that Medicare cost HMO enrollees are more like risk HMO enrollees than Medicare FFS beneficiaries, the inability to distinguish cost HMO from Medicare FFS hospital discharges in state hospital discharge data could mute some of the HMO-FFS differences in the California data.

2.4.2 HMO-FFS Differences in Demographic Composition

Table 2.3 contains information about the age-sex composition of the aged Medicare risk HMO and FFS populations in each of the study states. The relative patterns of FFS-HMO differences are fairly consistent among study states. Males comprise a little higher percentage of Medicare HMO risk enrollees than FFS beneficiaries in all of the states. Examination of the age distribution data by sex shows that this is probably due as much to the much lower prevalence of female HMO enrollees in the oldest age classes than to the higher prevalence of male HMO enrollees in any particular age class.

The age-sex distribution data clearly indicate the Medicare risk HMO population to be younger than its counterpart FFS population in all study states. A higher percentage of Part A person-years is uniformly found in the age classes below 75 years old for risk HMO enrollees than for FFS beneficiaries. The most notable differences in age composition are found in the "oldest old" age category of beneficiaries 85 years old and older. With the exception of males 85 years or older in Florida, a substantially higher percentage of aged Medicare FFS beneficiaries are found in this oldest age class.

Table 2.3: Age-Sex Beneficiary Population Distribution by State and HMO/FFS

	State of Residence									
Beneficiary Attributes	California		Florida		New York		Massachusetts		Total	
	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO
Part A person-years (in thousands)	2461	499	1934	262	2158	77	762	439	7315	882
Males										
% 65-69 yrs	15.3	16.1	14.6	14.9	14.3	21.0	13.8	17.0	14.7	16.2
% 70-74 yrs	11.3	12.2	11.3	12.7	10.4	10.8	10.4	12.1	10.4	12.2
% 75-79 yrs	7.6	8.0	8.1	9.2	7.2	6.5	7.2	7.6	7.6	8.2
% 80-84 yrs	4.3	4.2	4.8	5.2	4.2	3.3	4.1	3.6	4.1	4.4
% 85+ yrs	2.8	2.3	2.9	2.8	2.9	1.7	2.7	1.7	2.9	2.3
Females										
% 65-69 yrs	18.3	20.0	17.8	17.4	17.9	24.6	17.4	21.8	17.9	19.7
% 70-74 yrs	14.3	15.1	14.5	15.0	14.4	13.8	14.7	16.2	14.4	15.0
% 75-79 yrs	11.0	10.9	11.4	11.4	11.8	9.4	12.0	10.6	11.4	10.9
% 80-84 yrs	7.7	6.6	8.0	6.9	8.6	5.5	8.9	5.8	8.2	6.6
% 85+ yrs	7.3	4.7	6.6	4.3	8.3	3.4	8.9	3.7	7.6	4.4

2.4.3 HMO-FFS Differences in Hospital Use Rates

Table 2.4 contains empirical findings regarding hospital use rates for the study population as a whole and by HMO and FFS sectors in each state. The estimated hospital discharge and days of care rates for the combined aged populations of the four study states were 323 hospital discharges and 2,890 hospital days per 1,000 Part A person-years, respectively. Given that three of the study states have had higher than average Medicare hospital use rates historically, these estimates, which omit out-of-state hospital discharges, compare favorably with reported national rates of discharge and days of care for all beneficiaries of 316 hospital discharges and 2,642 hospital days per 1,000 beneficiaries in 1992 (Health Care Financing Administration 1995).

The top portion of Table 2.4 also contains estimated hospital discharge and days of care rates by age-sex groups and relative risk factors (defined as the age-sex group specific rate divided by the overall total aged population rate) for the pooled population from all study states. Again there is considerable face validity in the estimated rates. The estimated age-sex cell risk factors show the expected pattern of very low relative risk in the youngest age class of 65-69 years, and increasing risk of hospital use with age, particularly in the oldest age classes. Furthermore, the estimated hospital use rates follow the expected pattern of being higher in each age class for males than for females. These patterns compare favorably with trends in published national data for all Medicare beneficiaries (Health Care Financing Administration 1995).

The bottom portion of Table 2.4 contains age-sex adjusted hospital discharge and days of care rates for each of the study states broken down by HMO and FFS sectors. Both direct and indirect age-sex adjustments were applied to the data with no qualitative differences in the results given the rather modest overall age-sex composition differences reported earlier in Table 3. Table 2.4 contains the results of direct adjustments. The 1992 pooled age-sex distribution of Part A person-years for all aged beneficiaries in the four study states was employed in the computation of these directly adjusted rates.

Table 2. 4: Hospital Discharge and Hospital Day Rates by Age-Sex Class, State, and by HMO/FFS

Classification	Part A Person-years (in thousands)	Hospital discharge rate (/1,000)	Hospital days rate (/1,000)
All States	8199	323	2890
Males 65-69 yrs	1216	231 (0.72)	1783 (0.62)
70-74 yrs	909	323 (1.00)	2636 (0.91)
75-79 yrs	627	412 (1.27)	3580 (1.23)
80-84 yrs	359	514 (1.59)	4728 (1.64)
85+ yrs	230	631 (1.95)	6237 (2.16)
Females 65-69 yrs	1487	189 (0.58)	1529 (0.53)
70-74 yrs	1187	264 (0.82)	2267 (0.78)
75-79 yrs	933	334 (1.03)	3085 (1.07)
80-84 yrs	656	418 (1.29)	4136 (1.43)
85+ yrs	594	517 (1.60)	5374 (1.86)
California	2960	331	2463
HMO	499	189	1147
FFS	2461	359	2716
Florida	2196	300	2337
HMO	262	266	1669
FFS	1934	304	2421
New York	2235	324	3926
HMO	77	140	1455
FFS	2159	330	4002
Massachusetts	806	358	3053
HMO	44	272	1927
FFS	762	362	3104

At the state level, Massachusetts and Florida had the highest and lowest adjusted annual hospital discharge rates at the levels of 358 and 300 hospital discharges per 1,000 beneficiaries, respectively. While Florida also had the lowest annual days of care rate at a level of 2,337 hospital days per 1,000 beneficiaries, the highest days of care rate at the state level was found in New York at a level of 3,926 hospital days per 1,000 beneficiaries.

A comparison of HMO and FFS adjusted hospital discharge and days of care rates in Table 2.4 shows the expected finding of lower hospital discharge and days of care rates for risk HMO enrollees relative to their FFS counterparts in all study states. It is also apparent that the relative differences in HMO and FFS rates for annual hospital days is larger than for HMO-FFS differences in hospital discharge rates. For example, whereas the FFS hospital discharge rate of 362 discharges per 1,000 beneficiaries in Massachusetts was almost 33 percent higher than the rate of 272 discharges per 1,000 risk HMO enrollees in the state, the FFS days of care rate of 3,104 days per 1,000 beneficiaries was about 63 percent higher than the risk HMO rate of 1,927 days per 1,000 beneficiaries. Similar findings for all study states indicate that average lengths of stay for hospitalized Medicare risk HMO enrollees are uniformly lower than those of their counterpart FFS beneficiaries in all states.

While the estimated HMO utilization rates were lower than FFS rates in all of the study states, the relative HMO/FFS differences varied among the study states. In Florida and Massachusetts estimated FFS hospital discharge rates were only 14 and 33 percent higher than estimated HMO discharge rates. In contrast, the estimated HMO hospital utilization rates for both California and New York were quite low relative to the estimated FFS rates. Estimated FFS hospital discharge rates for New York and California were about 135 percent and 90 percent higher, respectively, than estimated HMO rates. Since extremely low Medicare risk HMO hospital discharge rates were estimated for a number of subareas in these states, it is possible that risk HMO hospitalizations may be undercounted in these states unless these differentials are corroborated by very large relative HMO/FFS differences in

estimated mortality rates in these states as well. Undercounts of Medicare risk HMO hospitalizations would have obvious adverse implications toward analysis of geographic variations in levels of hospital utilization for HMO versus FFS beneficiaries. However, this should not seriously bias the comparisons of the diagnostic composition of FFS versus HMO hospitalizations given the much larger volume of FFS hospitalizations unless possible undercounting is the result of a systematic process for hospitalizations with certain diagnostic codes. Vendors of the hospital discharge were unaware of any systematic coding problems with the expected payor field.[3]

2.4.4 HMO-FFS Differences in Age-Sex Adjusted Mortality Rates

Table 2.5 contains empirical findings regarding unadjusted and age-sex adjusted mortality for the study population as a whole and by HMO and FFS sectors in each state. Unadjusted death rates were computed from the HCFA Denominator file data by dividing the number of beneficiary deaths occurring during the 1992 calendar year by the number of beneficiaries alive at the beginning of the year. Surviving beneficiaries were assigned to risk HMO or FFS sectors on the basis of their HMO enrollment status in December, 1992. Deaths were assigned to the risk HMO or FFS sector on the basis of risk HMO enrollment status in the month of death.

---

[3] Undercounting of HMO hospitalizations could result from the incorrect assignment of Medicare as the primary payor for elder patients from risk HMOs. It could also result if certain acute care hospitals, e.g., those affiliated or owned by HMOs, do not submit hospital discharge data. To date, the only primary payor coding problems associated with specific hospitals potentially contributing to undercounts of HMO hospitalizations for which we could obtain information about were for several hospitals serving patients in central Massachusetts. The Massachusetts hospitalization rates reported on Table 4 reflect these potential HMO undercounts. In the case of New York and California it was acknowledged that the primary payor field on discharge records is likely to understate Medicare HMO hospitalizations. The payor field on discharge records is strictly defined as the expected source of payment at discharge and is not used for direct billing purposes. Given the dominance of Medicare over HMO as the primary payor for elders, it is more likely for an elder Medicare HMO enrollee to be incorrectly coded as Medicare, than for a elder FFS beneficiary to be incorrectly coded as HMO.

Table2. 5: Unadjusted and Age-Sex Adjusted Annual Death Rates by State, and by HMO/FFS

Classification	Number of Deaths in 1992	Unadjusted death rate	Age-sex adjusted death rate
State			
California	142247	4.39%	4.51%
FFS	126975	4.75%	4.79%
HMO	15272	2.70%	3.04%
Florida	98054	4.25%	4.33%
FFS	87251	4.32%	4.36%
HMO	10803	3.81%	4.14%
New York	114341	4.73%	4.65%
FFS	112263	4.81%	4.65%
HMO	2078	2.43%	4.79%
Massachusetts	40274	4.75%	4.61%
FFS	38366	4.79%	4.59%
HMO	1908	4.04%	4.89%
All States	394916	4.48%	4.51%
FFS	364855	4.66%	4.63%
HMO	30061	3.06%	3.46%

Since new HMO enrollees and HMO disenrollees in 1992 will not be at risk of death for a full year within their assigned end of year HMO/FFS sector, the estimated simple death rates have the shortcoming that Medicare risk HMO mortality rates will vary depending upon both the relative volume of and degree of selectivity in new HMO enrollment and disenrollment within the year. In general Medicare risk HMO death rates will be biased downward due to "truncated calendar year" effects, the greater is the volume of new enrollment to disenrollment, the greater is new enrollment relative to full-year HMO enrollment, and the greater is the degree of HMO favorable selection in new enrollment and disenrollment.[4] While comparisons of the ratios of Part A eligible person-years to beginning-of-year counts of beneficiaries for Medicare risk HMO and FFS sectors show no large differences suggesting truncation effects may be modest, the selectivity of new HMO enrollment and disenrollment could not be measured independently.

Age-sex adjustments to simple death rates were made using both direct and indirect adjustment methods. The adjusted rates reported in Table 2.5 were computed with indirect methods on the basis of 1992 national annual death rates for the same 10 age-sex population groups used for adjusting hospital use rates (National Center for Health Statistics 1993).

Examining the 1992 state death rates for the combined HMO and FFS Medicare beneficiary populations, the differences among states are fairly modest, particularly after adjustments were made for differences in the age-sex population composition among states. The adjusted mortality rates for New York (4.65%) and Massachusetts (4.61%) were only a little higher than those of Florida (4.33%) and California (4.51%). While these mortality rate differences are more modest than differences in

---

[4] While the net effects of a truncated year are complex, it should be obvious that if a Medicare beneficiary enrolls in an HMO in the middle of the year they must have survived one half of a year before HMO enrollment. Whereas a full-year HMO enrollee is at risk of death during the first half of the year, the new mid-year HMO enrollee is not at risk of death from the perspective of the HMO population in the first half of the year.

hospital use rates reported in Table 2.4, the rank order of adjusted mortality rates among study states does correspond to the rank order of adjusted days of care rates presented earlier.

With the exception of Massachusetts where the adjusted Medicare HMO mortality rate actually exceeded the FFS rate, HMO-FFS differences in adjusted mortality rates conformed to a priori expectations. In the case of Florida, the adjusted FFS mortality rate was about only about 5 percent higher than the Medicare risk HMO rate. The magnitude of this difference is comparable to the 14 percent higher FFS hospital use rate reported earlier. The HMO-FFS differences in adjusted mortality rates in California and New York were quite large. The adjusted FFS mortality rates in California and New York were about 58 percent and 52 percent higher than Medicare risk HMO mortality rates, respectively. While these differences are large, it is not clear that they can fully account for the even larger estimated differences in hospital use rates reported earlier in Table 2.3. Finally, it is interesting that the adjusted mortality rate for Medicare risk HMO enrollees in Massachusetts was actually higher than that of their FFS counterparts. To the extent that selective new HMO enrollment offsets the effects of an aging HMO enrollee population, the relatively high Medicare HMO death rate in Massachusetts may reflect the stagnant level of Medicare risk HMO enrollment in the state during the late 1980s and early 1990s.

2.4.5 HMO-FFS Differences Diagnostic Composition of Hospitalizations

The descriptive data examined thus far show a varied set of HMO and FFS populations with varying hospital use rates in the study states. At this point we will examine the diagnostic composition of the hospitalizations in terms of their DCG risk and discretion class distributions.

DCG Risk Class Distribution

Table 2.6 contains the distribution of hospital discharges and days of care among DCG risk classes by Medicare risk HMO and FFS sectors for each of the study states. Examination of Table 2.6 reveals only very marginal differences in the distribution of discharges and days of care among DCG risk classes for Medicare risk HMO and FFS beneficiaries both within and among states. While the observed differences are significant at conventional levels of statistical significance given the large sample sizes involved, it would be difficult to conclude that there is much practical significance in the differences given the much larger differences in hospital use rates observed. Furthermore, there are no obvious noteworthy systematic patterns among the distributions with respect to overall hospital use rates or population mortality rates displayed earlier. The relatively invariant DCG class distributions were not the result of aggregation over the original eight risk DCG classes of Ellis and Ash (1995). Similar invariant patterns were found for the disaggregated eight DCG risk classes as well, but are not reported here.

Overall, the data on Table 2.6 suggest that, at least in aggregate, there is very little difference between Medicare risk HMO and FFS distributions of hospitalizations among relative DCG risk classes. Since high discretion hospitalizations are assigned to the lowest DCG 0 risk class, the relative invariance of the distribution suggests reductions in hospital use associated with lower HMO hospital use rates are not likely to be due to reductions in highly discretionary hospital use, at least as defined under the DCG model discretion ratings.

Discretion Class Distribution

To the extent that the DCG discretion scores are valid measures of relative discretion, the distribution of hospital discharges and days of care among discretion classes provides a direct means of testing the hypothesis that lower hospital use rates in Medicare risk HMOs relative to Medicare FFS

Table 2.6: DCG Distribution Among Hospital Discharges and Days of Care by HMO/FFS and by State

	State of Residence									
	California		Florida		New York		Massachusetts		Total	
Percent by DCG Class	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO
Percent of Discharges										
% DCG 0	68.7%	68.7%	68.3%	66.6%	67.1%	67.6%	67.0%	67.1%	67.9%	67.8 %
% DCG 1-2	11.3%	12.1%	11.4%	12.7%	12.0%	13.9%	12.7%	14.4%	11.7%	12.5 %
% DCG 3-4	9.9%	9.9%	9.4%	8.7%	9.6%	9.6%	8.8%	8.0%	8.0%	8.9%
% DCG 5-7	10.1%	10.2%	70.9%	12.0%	12.0%	8.9%	11.5%	9.5%	10.8%	10.8 %
Percent of Hospital Days										
% DCG 0	70.4%	71.7%	70.9%	71.1%	69.7%	70.9%	70.6%	72.0%	70.2%	71.4 %
% DCG 1-2	9.1%	8.7%	9.4%	10.0%	9.6%	10.5%	9.7%	10.6%	9.2%	9.6%
% DCG 3-4	11.4%	10.7%	9.4%	8.2%	10.1%	9.6%	8.3%	8.0%	10.2%	9.5%
% DCG 5-7	9.1%	8.0%	10.0%	10.7%	11.2%	9.2%	11.3%	9.4%	10.4%	9.6%
Total discharges (in thousands)	880	91	590	69	718	10	279	12	2468	182
Hospital discharge rate (/1,000)	359	189	304	266	330	140	362	272	338	221

use rates are attributable to HMO reductions in discretionary hospital use. If lower Medicare HMO use rates are largely the result of reductions in discretionary hospital use (as defined by DCG discretion scores) then HMOs should have disproportionately fewer high discretion hospitalizations and days of care and disproportionately more low discretion hospitalizations than the Medicare FFS sector.

Table 2.7 contains the distribution of hospital discharges and days of care among high, moderate, and low discretion classes by Medicare risk HMO and FFS sectors for each of the study states. These data provide little empirical support for the main study hypothesis.[5] The top portion of the table contains the HMO and FFS discretion class distributions for hospital discharges. While Medicare risk HMOs had slightly smaller percentages of high discretion hospitalizations in California and New York, the findings for Massachusetts and Florida contradict the main study hypothesis. In both of these states a higher percentage of HMO than FFS hospital discharges were classified as high discretion. With the exception of Massachusetts, there were a higher percentage of low discretion hospitalizations among risk HMO discharges. The practical HMO-FFS differences were only marginal, however.

The empirical findings for the distribution of hospital days among discretion score categories in the bottom of the table were consistent with the main study hypothesis. A smaller share of Medicare risk HMO hospital days were associated with high discretion hospitalizations and hospital days for low discretion hospitalizations comprised a larger share of risk HMO hospital days. While the differences are statistically significant due to the large sample size, the differences are much too small to account for more than a minuscule portion of the HMO-FFS differences in hospital use rates even in the states

---

[5] Again the results do not result from aggregation over the original DCG discretion scores of Ellis and Ash which ranged from 3 to 12. Similar invariance in diagnostic composition was found for those 10 diagnostic groups defined by discretion score.

Table 2.7: Distribution of Hospital Discharges and Days of Care among Discretion Classes by HMO/FFS and State

	State of Residence									
	California		Florida		New York		Massachusetts		Total	
Percent by Discretion Class	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**
Percent of Discharges										
% Low Discretion (3-4)	28.1	25.9	26.4	27.7	27.6	27.7	27.1	25.9	27.5	28.5
% Moderate Discretion (5-8)	33.0	34.5	34.7	32.5	32.5	32.5	32.7	33.3	33.2	32.3
% High Discretion (9-12)	38.9	36.0	38.9	39.9	39.8	39.9	40.2	40.8	39.3	39.2
Percent of Hospital Days										
% Low Discretion (3-4)	32.1	36.0	30.9	34.1	34.1	36.7	33.4	33.3	32.8	35.2
% Moderate Discretion (5-8)	32.3	30.0	34.1	31.7	30.1	28.5	31.3	31.2	31.7	30.6
% High Discretion (9-12)	38.9	34.5	35.1	34.2	35.8	34.8	39.3	34.5	35.5	39.2
Total discharges (in thousands)	880	91	590	69	718	10	279	12	2468	182
Hospital discharge rate (/1,000)	359	189	304	266	330	140	362	272	338	221

of Massachusetts and Florida where the higher estimated risk HMO hospital rates are more likely to be valid.

2.5 DISCUSSION

This descriptive analysis of the diagnostic composition of Medicare HMO and FFS hospitalizations has provided little empirical support for the main study hypothesis, namely, that a significant portion of lower Medicare risk HMO hospital admission rates is associated with HMO success in reducing discretionary hospitalizations. Our descriptive analysis of over 2.6 million hospital discharges in four states shows the diagnostic composition of Medicare risk HMO and FFS hospitalizations to be relatively invariant with respect to both DCG risk class and discretion score. That is, both Medicare HMO and FFS hospitalizations have very similar distributions of hospitalizations among DCG classes and among discretion score classes. This result was found for each of the four study states. Under the DCG risk cell and discretion score diagnostic classifications at least, it does not appear that Medicare risk HMOs disproportionately reduce certain kinds of hospital admissions over others. Our findings are consistent with Luft's (1978) findings of across the board lower HMO use rates for both discretionary and nondiscretionary hospitalizations. The current study expands on Luft's work by the employment of explicit physician ratings in defining discretionary hospitalizations and by use of much larger population samples.

The relative invariance of the diagnostic composition of Medicare risk HMO and FFS hospitalizations with respect to discretion level may be consistent with three alternative interpretations offered by Luft (1978) for similar findings. First, the DCG discretionary scores derived from physician judgments for diagnostic codes may simply be inadequate measures of discretion. There may "discretionary" hospitalizations within any diagnostic category of hospitalizations depending upon how discretion is defined. Second, unspecified HMO-FFS health status differences may result in much

lower non-discretionary hospital use by Medicare risk HMO enrollees that counterbalance HMO reductions in discretionary service use. Given the large differences in age-sex adjusted HMO-FFS mortality rates for two of the states this explanation cannot be ruled out entirely. However, this explanation is inconsistent with our findings of similar diagnostic patterns for the two states where HMO-FFS mortality rate differences were much smaller. Thirdly, it is possible that Medicare risk HMOs systematically undertreat and Medicare FFS providers overtreat nondiscretionary conditions. While further analyses of patient outcomes are needed to fairly test this third potential explanation offered by Luft (1978) for similar findings, our comparison of adjusted HMO and FFS mortality rate outcomes do not support this explanation. That is, with the exception of Massachusetts, age-sex adjusted mortality rates were lower for Medicare risk HMO enrollees relative to their FFS counterparts.

The most feasible of the three alternative explanations for this study's findings may be that DCG physician discretion ratings are simply flawed in the sense that the ratings simply may not distinguish hospitalizations which are highly discretionary very well. In order to sort out whether the empirical findings in this chapter are attributed to problems in the DCG discretionary ratings for specific diagnoses, a broader problem of rating the discretion level of hospitalizations, or possibly the absence of measurable differences in HMO and FFS practice styles, additional analyses are needed. Subsequent chapters will report analyses which further address the question of the validity of the DCG discretion measures through:

- comparisons of the overlap of DCG discretion score categorizations of hospital discharges with several alternative definitions of discretionary hospitalizations in the literature;
- comparisons of HMO-FFS differences and geographic differences in diagnostic composition of Medicare hospitalizations for these alternative discretion measures.

The findings from those analyses will provide additional objective empirical data for assessing whether the discretionary component of the DCG model of Ellis and Ash (1995) serves as a useful means of reducing potential biases in risk classification associated with differences in medical practice styles.

A HMO AND FFS DISCRETIONARY HOSPITAL ADMISSIONS: A COMPARISON OF ALTERNATIVE DISCRETION CLASSIFICATIONS

CHAPTER 3

3.1 INTRODUCTION

In the previous chapter we compared the distribution of hospital discharges and hospital days among DCG discretion score categories for Medicare FFS beneficiaries and Medicare risk HMO enrollees. The data showed only very modest differences in the diagnostic composition of hospitalizations by broad DCG discretion class between HMO and FFS sectors and among the four study states. While caution must be exercised in drawing conclusions from simple descriptive analyses of data from a sample of four states, the analyses do not provide empirical evidence that is supportive of the validity of the physician discretion component of the DCG model of Ellis and Ash (1995). It is clear that the DCG model discretion ratings do not distinguish among hospitalizations in such a way as to account for very much of observed HMO-FFS and interstate differences in hospital use rates for the aged Medicare population.

In this chapter we will investigate whether employment of alternative classifications of discretionary hospital admissions in the health services research literature yields produce more supportive empirical findings to those found for the DCG discretion model. In the following section we review the development of alternative definitions of discretionary hospitalizations in the research of Anderson, et al. (1989) and Roos, et al. (1988). This is followed in the third section by an empirical analysis of the agreement between the assignments of hospital discharges among high, moderate, and low discretion categories under the DCG model and three alternative discretion classification systems derived from the work of Anderson, et al (1989) and Roos, et al. (1988). Empirical findings about state-level HMO-FFS differences in the discretionary composition of hospital discharges and days of care for the three alternative classification systems are presented and compared with the DCG

discretion score findings from the last chapter in the fourth section. The last section contains a summary of the findings and closing remarks.

3.2 ALTERNATIVE DEFINITIONS OF DISCRETIONARY HOSPITALIZATIONS

3.2.1 Patient Variation and Physician Discretion Classifications of Anderson, et al. (1989)

While this research is focused on the physician discretion ratings of the DCG model of Ellis and Ash (1995), Anderson, et al. (1986) were actually the first to develop a prior use risk adjustment model where level of physician discretion in the decision to hospitalize was taken account of. Several physicians rated all three-digit ICD-8 codes into one of three "discretion" categories (low, medium, high) based on the extent to which it was believed that there was room for physician discretion in the decision to hospitalize. Prior use models for predicting future Medicare reimbursement levels were estimated in which high discretion hospitalizations were excluded from as well as included in prior hospital use measures. Models in which high discretion hospitalizations were excluded were shown to perform nearly as well as those in which physician discretion was not distinguished.

Anderson, et al. (1989) extended this preliminary research through a larger survey of 169 physicians from 31 specialties in both academic and FFS settings regarding their judgements concerning physician discretion for various diagnostic codes and procedures. Physicians were asked to rate diagnoses and procedures through separate questionnaires related to their specialty. The questionnaire of interest to the proposed study was the diagnosis questionnaire which included questions concerning chronicity, a patient variation component of discretion, physician discretion, probability of future readmission, and expected future expense level.

Anderson, et al.'s (1989) concept of discretion was actually comprised of two separate dimensions: patient variation and physician discretion. The "patient variation" dimension of discretion

was addressed by asking physicians about the likelihood (high, medium, low) that an average physician confronted with a patient with a given primary diagnosis would admit the patient to a hospital. This aspect of discretion reflects the expected severity of patients seen by average physicians. High discretion is equated with an assessment of a low probability of admission for patients with a given medical condition. If there is greater variability in severity level of patients with a given medical condition there should be a lower likelihood that any individual patient would be hospitalized by the average physician. Low discretion on this component of discretion is equated with a high probability of hospital admission. If there is little variability in the severity of patients with a given medical condition and hospitalization is appropriate form of treatment, there should be a high probability that an average physician will hospitalize all such patients. Table 3.1 contains a sample of diagnoses assigned to high low, moderate, and high patient variation categories.

The "physician discretion" component of the discretion of Anderson, et al. (1989) was intended to be similar in concept to the DCG physician discretion ratings developed by Ellis and Ash (1995). It was addressed through questions about the degree to which different physicians might vary in deciding whether to admit the same patient with a given primary diagnosis. Physicians were asked to rate each diagnosis as one with high, medium, or low variation among physicians regarding the decision to admit. Examples of common diagnoses were included in each discretion category to provide physicians with a reference point for making their own ratings.

Analyses of the physician discretion ratings by Anderson, et al. (1989) revealed some moderate but statistically significant differences in the physician discretion ratings by HMO and academic FFS physician survey respondents. While both physician groups rated 7 percent of the 200 most common diagnoses as involving high physician discretion, HMO physicians rated fewer diagnoses as low discretion (58%) versus academic physicians (67%). After employing some rules for classifying diagnoses where there were inconsistent ratings among physicians, each diagnosis was assigned to a

Table 3.1: Anderson Patient Variation Ratings for Selected Common Medicare Conditions

Discretion Category	3-4 digit ICD-9 code
Low Discretion	
Fracture of Tibia and Fibula	823
Acute Myocardial Infarction	410
Gastrointestinal Hemmorage	578
Malignant Neoplasm colon	153
Essential Hypertension	401
Moderate Discretion	
Lymphoproliferative Disease, NOS	238.7
Meningitis	340
Phlebitis and Thrombophlebitis	451.1
Duodenal Ulcer	532
Pneumonia Organism NOA	486
Emphysema, NOS	492.8
Heart Failure	428
High Discretion	
Cataract	366
Asthma	493
Nasal Polyp, NOS	471.9
Bunion	727.1

high, moderate, or low physician discretion class. Table 3.2 contains a sample of common diagnoses and their assignment to physician discretion categories.

Two of the more surprising findings from physician survey ratings was the small proportion of diagnoses rated as having high levels of physician discretion, and the marked differences with the physician discretion ratings of Ellis and Ash (1995). Anderson, et al. (1989) reported that for the 199 ICD-9 codes with greatest frequency as a principal diagnosis for aged Medicare inpatient hospitalizations, there was exact agreement with the DCG discretion ratings in only 35 percent of the codes and complete disagreement with opposite ratings (high versus low discretion) for 24 percent of the codes.

3.2.2. Index of Discretion Classification of Roos, et al. (1988)

Under the "professional uncertainty" thesis advanced by Wennberg, et al. (1982), it is maintained that there is physician discretion toward hospitalization and/or the performance of certain surgical procedures because the relative clinical outcomes of treatment alternatives may not be well established. Since physicians may disagree among themselves concerning the most appropriate course of care for a given medical condition, the decision to hospitalize may be highly discretionary. Given this lack of consensus about what is the appropriate treatment, local physician practice styles towards the use of hospital treatment may be influenced not only by general physician attitudes, but also by factors such as the availability of hospital beds (Wennberg, et al. 1987).

Under this thesis it can be deduced that the degree of geographic variations in per capita hospital admission rates for given medical conditions should tend to vary inversely with the degree of professional consensus about the appropriateness of acute hospitalization as a course of treatment. Indeed, lesser geographic variations have been found for conditions such as myocardial infarction and inguinal hernia repair where current medical standards generally dictate inpatient hospital treatment.

Table 3.2:Anderson, et al. (1989) Physician Discretion Ratings for Selected Common Medicare conditions

Discretion Category	3-4 digit ICD-9 code
Low Discretion	
Fracture of Tibia and Fibula	823
Acute Myocardial Infarction	410
Nasal Polyp, NOS	471.9
Phlebitis and Thrombophlebitis	451.1
Bunion	727.1
Moderate Discretion	
Lymphoproliferative Disease, NOS	238.7
Meningitis	340
Duodenal Ulcer	532
Pneumococcal Pneumonia	481
Emphysema, NOS	492.8
High Discretion	
Hyperplasia of Prostrate	600
Viral Pneumonia	480
Essential Hypertension	401
Arteriosclerosis	440
Cataract	366

Hospital admission rates for certain surgical procedures where there has been controversy about their value, such as hysterectomy, prostatectomy, and tonsillectomy, have been shown to exhibit more substantial geographic variations (e.g., Wennberg and Gittelsohn 1982; McPherson, et al. 1982; Roos, et al. 1988).

The recent study of Roos, et al. (1988) classified hospitalizations into an "index of discretionary admissions" on the basis of consistent patterns in the relative level of geographic variability in hospital discharge rates in states of Iowa, California, Massachusetts, and Maine. Roos, et al. (1988) defined 77 closely related modified surgical and medical DRGs by aggregating lower volume DRGs on the basis of same primary diagnosis, related operations for some surgical DRGs, and major disease category. Variations in age-sex adjusted hospital discharge rates among hospital service areas in each state were compared using systematic coefficients of variation (SCVs). Modified DRGs were ranked by SCV and stratified into deciles for each state. A consistent pattern of variation was said to exist among states when decile rankings were within three deciles in at least three of the four study states. A five point index of discretionary admissions (low variation, moderate variation, high variation, very high I variation, very high II variation) was developed from the SCV scores. Table 3.3 contains examples of the discretionary classification for selected modified DRGs.

In contrast to the discretion ratings of Ellis and Ash (1995) and Anderson et al. (1989) which were based on a priori physician clinical judgements, the discretion index of Roos et. al (1988) was empirically derived based on consistent patterns of geographic variations. Because of the manner in which these index of discretion ratings were developed, they are expected to better distinguish among subgroups of hospitalizations with higher versus lower geographic variability in discharge rates than the clinically-based DCG and Anderson discretion classification systems. Whether this will result in superior performance in distinguishing hospitalizations which largely account for observed geographic and HMO-FFS variations in hospital use rates is an empirical question.

Table 3.3: Roos, et al. (1988) Index of Discretion Ratings for Modified DRGs

Discretion Category	DRGs
Low Variation	
Acute Myocardial Infarction	121-123
Specific Cerebrovascular Disorders	14
Gallbladder Disease with Cholecystecomy	195-198
Inguinal and Femal Hernia	161-162
Infections of Kidney	590
Moderate Variation	
Gastrointestinal Hemmorage	174-175
Appendicitis with Appendectomy	164- 167
High Variation	
Angina Pectoris	140
Gastrointestinal Obstruction	180-181
Respiratory Neoplasms	82
Heart Failure and Shock	127
Transurethal Operations	310-311
Very High Variation I	
Kidney and Urinary Tract Infections	320-321
Cellulitis	277-279
Adult Gastroenteritus	182-183
Red Blood Cell Disorders	395-396
Back and Neck Operations	214-215
Very High Variation II	
Chronic Obstructive Lung Disease	88
Tonsillectomy	59-60
Hypertension	134
Adult Bronchitis and Asthma	96-97
Extraocular Operations	40-41

3.3 DATA AND METHODOLOGY

3.3.1 Data Sources

The main source of data for the analysis were the public use hospital discharge files for California, Florida, Massachusetts, and New York discussed in Chapter 2. Hospital discharges were assigned to categories of the Anderson patient variation and physician discretion classifications through ICD-9 CM codes for the principal diagnosis of hospital discharges. With the exception of Massachusetts, the grouper version 9 of federal DRG codes was used to classify hospital discharges under the Roos index of discretion. In Massachusetts, the 1992 calendar year hospital discharge file had to be constructed from two raw hospital discharge data files that were organized by fiscal year. The later fiscal year file contained grouper version 10 DRG codes. We could not directly assess the impact of changes between the two DRG grouper versions on hospital discharge classifications in Massachusetts. However, the New York hospital discharge file contained DRG codes for both DRG grouper versions. A comparison of the distributions of DRG codes for hospital discharges in New York based on these two groupers revealed only marginal differences in DRG code assignments.

3.3.2 Impacts of ICD-9 CM Coding Changes

The Roos index of discretion was developed with state hospital discharge data files from the years 1980 through 1983 where discharges were assigned to DRGs through DRG grouper version 3. The physician ratings employed by Anderson, et al. (1989) were based on interpretations of ICD-9 CM codes as of 1986. Since assignments to discretion and/or variation categories in this study had to be done with 1992 ICD-9 codes, it was important to investigate whether changes in the interpretation of ICD-9 codes over the intervening years would likely result in significant changes in the intended classifications of 1992 study hospital discharges. In Chapter 2 it was noted that a clinical assessment of

changes in interpretation of ICD-9 codes on DCG risk class and discretion score classification suggested that any impacts were negligible. Similar analyses were conducted for the alternative discretion classification systems of Anderson et. al (1989) and Roos, et. al (1988).

A file containing a direct mapping of 4 and 5 digit ICD-9 codes to high, moderate and low categories of the "patient variation" and "physician discretion" classifications of Anderson et. al (1989) were obtained from the main author.[6] Frequency distributions of hospital discharges from the combined pool of data for the four study states were compiled and sorted in terms of the relative frequency of the joint classification of the "patient variation" and "physician discretion" ratings of Anderson et. al (1989). Frequency distributions for the five-digit ICD-9 CM codes of the principal diagnoses within each of these joint categories were then compiled and sorted by relative frequency. A physician clinical consultant compared the more frequent principal diagnoses in each joint "patient variation-physician discretion" classification with a cumulative record of ICD-9 CM code interpretation changes occurring between 1986 and 1992 compiled by the Health Care Financing Administration to detect situations where changes in coding interpretations might affect the classification of hospital discharges. No situations were detected where code interpretation changes would have an impact of much significance.

The publication of Roos et. al (1988) reported the specific DRG codes for each modified-DRG group which was assigned to an "index of discretion" category. Communication with the main author of the study confirmed that DRG code assignments were made with DRG grouper version 3 and that all DRG codes not listed in the published article exhibited inconsistent patterns of geographic variability. The potential impacts of ICD-9 CM coding changes on the classification of hospital discharges among

---

We originally intended to incorporate the earlier discretionary ratings of Anderson, et al. (1986) in the study as well. Gerald Anderson informed us that the earlier study was a pilot study for their later research and that a complete mapping of ICD-8 codes to discretion categories no longer existed.

the Roos index of discretion categories were evaluated in a manner similar to that described above for the Anderson, et al. (1989) discretion classifications. First, frequency distributions of hospital discharges for all DRG codes were compiled with pooled discharge data from the four study states and sorted in terms of relative frequency. Within each DRG code group, frequency distributions of five-digit ICD-9 CM codes were compiled and frequent codes compared with a cumulative history of interpretation changes in ICD-9 codes between 1986 and 1992 by a consulting physician. No changes were found which were likely to have a significant impact on Roos index of discretion classifications.

Since the DRG assignments in the study data of Roos et. al (1988) were made with the DRG grouper version 3 and DRG assignments in our study data were made with the DRG grouper version 9, some of these DRG assignments could differ due to application of different groupers. We sought to assess the likely impacts of these changes upon the categorical assignments of hospital discharges. Unfortunately, we were unable to obtain the necessary code for DRG grouper version 3 which would permit us to estimate the magnitude of changes in DRG assignments. While significant changes in the distribution of hospital discharges among the five categories of the index of discretion of Roos et. al (1988) are unlikely to result from grouper changes, the reader is cautioned that no data were evaluated to support this assertion.

In addition to assessing potential impacts of ICD-9 CM code interpretation changes on classification of hospital discharges, the consulting physician reviewed the higher frequency principal diagnoses from the pooled study file of hospital discharges for situations where changes in medical practice styles over time, possibly associated with technological developments or changed reimbursement practices, might affect physicians' ratings of "discretion." Given the much smaller number of DRGs involved in the Roos index of discretion, all DRGs assigned to any of the five level of variation categories were reviewed regardless of the relative frequency of hospital discharges.

The most notable case where significant changes in medical practice style were likely to affect judgements about discretion was the "extraocular operations" modified DRG of Roos et. al (1988) which was comprised of DRGs 40-41. This modified DRG includes the 3-digit ICD-9 code 366 for cataracts in the DCG discretion classifications. Whereas there may have been considerable discretion in physician decisions' to hospitalize patients for cataract surgery in the early 1980s which contributed to the classification of DRGs 40-41 in the highest variation category (and ICD-9 366 to the high discretion category in the DCG risk classification as well), there was probably very little physician discretion involved in hospitalizing Medicare patients for cataract surgery under DRGs 40-41 in 1992 with the widespread performance of outpatient cataract surgery. Since there were only 1,216 hospital discharges in the entire pooled discharge file of over 2.6 million discharges with DRG 40-41, this change in medical practice style should not have much impact on the empirical findings. No change was made to the original Roos et. al (1988) classification of DRG 40-41 or to DCG and Anderson classifications of ICD-9 code 366. For similar reasons, no changes were made to the original Roos et. al (1988) classifications for DRGs which would be rarely ever, if at all, found in an aged Medicare population regardless of medical practice style issues.

3.3.3 Methodology

The two main research questions addressed in this chapter are:

- **How similar are the assignments of hospital discharges to DCG discretion classifications with assignments based on the alternative discretion classifications of Anderson et. al (1989) and Roos et. al (1988)?**
- **To what extent are HMO-FFS differences in hospital utilization rates associated with higher/lower rates of high discretion or high variation hospital discharges as defined under the classification systems of Anderson et. al (1989) and Roos et. al (1988)?**

A common set of categories were defined to compare DCG discretion levels with those of the three alternative discretionary classification systems. The Anderson patient variation and physician

discretion ratings each have three categories which correspond to high, moderate and low levels of discretion. As already discussed in the previous chapter, the DCG discretionary score ratings ranging from 3 to 12 can be meaningfully collapsed into high, moderate, and low discretion categories as well. Comparisons of the DCG discretion classifications with the five-category Roos index of discretion are complicated by two factors. First, there are numerous hospital discharges with DRGS that were not classified in any of the five index categories because of inconsistent patterns of geographic variation. Second, while the five categories are ordinal with respect to level of variation, the labels applied to the index categories (i.e., low variation, moderate variation, high variation, very high variation I, and very high variation II) are not balanced among high, moderate, and low variation categories.

A case could be made for aggregating the top three " high variation" categories into a single high variation category if an objective definition of "high variation" existed. However, any definition of high variation will be arbitrary. Furthermore, given the ambiguity associated with what should be done with unclassified hospital discharges, a better case can be made for assigning the highest variation categories I and II as "high variation," the lowest variation category as "low variation," and all other discharges, including unclassified ones, into a residual "moderate variation" category. Under this categorization the residual category is technically comprised of DRGs with geographic variations that are neither very high or very low. Since inconsistent geographic patterns of high and low variation are likely to be moderate on average, it is appropriate to assign DRGs with inconsistent variation to a residual middle category. This latter aggregation strategy was chosen for the creation an exhaustive three-category index of discretion from the original DRG classifications of Roos et. al (1988). Use of this aggregated index facilitates comparisons of the Roos index of discretion with the other three-category classification systems.

3.4 EMPIRICAL RESULTS

3.4.1 Agreement Between DCG Discretion and Alternative Classification Systems

Anderson Patient Variation Classifications

Table 3.4 contains a cross-classification of some 2.6 million hospital discharges for the study states with respect to the DCG discretion categories of Ellis and Ash (1995) and the Anderson patient variation discretion measure. The original low, moderate, and high likelihood of admission ratings of "patient variation" reported in Anderson et. al (1989) have been inversely recoded into low, moderate, and high discretion categories in the table to facilitate the comparison of the two classification systems. Under this inverse recoding scheme, hospital discharges originally classified as ones where patients have a "high probability" of admission by an average physician are recoded as "low discretion" hospitalizations. Hospital discharges of patients with diagnoses where there was a "low probability" of admission by an average physician are recoded as "high discretion" hospitalizations. Under this recoding scheme, the diagonal cells of the table represent counts of hospital discharges where there is conceptual agreement between the DCG discretion and patient variation measures.

Examining the cells of Table 3.4 and the corresponding cell/total percentages, the highest level of agreement was found for hospital discharges classified as low discretion under the DCG physician ratings. Nearly three-quarters of hospital discharges classified as low discretion in the DCG model were also classified as low discretion on patient variation, as reflected in a high probability of admission. There was much lower agreement between the two classification systems for hospitalizations classified as high discretion under the DCG model. Less than 30 percent of hospital discharges classified as high discretion under the DCG model were also classified as high discretion on patient variation with a low probability of admission. Overall, there was agreement in assignment for 43 percent of all hospital discharges. Complete disagreement in assignment, as reflected in opposite assignments (i.e., high with low), was found for about 10 percent of discharges.

Table 3.4: Cross-Classification of DCG Discretion Score Classifications with Anderson Patient Variation Classification

	Anderson Patient Variation Class				
DCG Discretion Class	Low Discretion	Moderate Discretion	High Discretion	Total Discharges	Marginal Distribution
Low Discretion	514.75 (19.42%)[1]	212.24 (8.01%)	2.44 (0.09%)	729.42	27.52%
Moderate Discretion	422.79 (15.95%)	326.68 (12.33%)	129.34 (4.88%)	878.80	33.16%
High Discretion	254.61 (9.61%)	490.27 (18.50%)	297.18 (11.21%)	1042.0	39.32%
Total Discharges (thousands)	1192.14	1029.18	428.96	2650.3	
Marginal Distribution	44.98%	38.83%	16.19%		100%

Percent of Agreement: 42.96%
Percent of Complete Disagreement: 9.70%

[1] Cell percentages of total discharges are reported in parentheses.

Anderson Physician Discretion Classifications

Table 3.5 contains the cross-classification of hospital discharges by DCG discretion class and Anderson physician discretion class. Overall there was agreement in the two classification systems for a little more than one-third of all hospital discharges, and there was complete disagreement for about 13 percent of hospital discharges. Similar to the findings for the Anderson patient variation classification, the greatest level of agreement was for hospitalizations classified as low discretion under the DCG model. Nearly two-thirds of low discretion hospitalizations under the DCG model were also rated as low discretion under the Anderson physician discretion ratings. There was least agreement in the classification of high discretion hospitalizations. Less than 2 percent of high discretion hospital discharges under the DCG model were similarly classified as high discretion under the Anderson physician variation ratings.

Anderson et. al (1989) noted that at a conceptual level, their physician discretion ratings are more similar to the DCG discretion ratings than are their patient variation ratings. It is interesting that a lower level of agreement was found between the physician discretion ratings and the DCG discretion ratings than for the patient variation ratings reported earlier in Table 3.4. It suggests that the DCG discretion ratings regarding substitutability of outpatient for inpatient care may have been highly influenced by physician rater perceptions about variations in the severity levels of patients.

Roos Index of Discretion Classifications

Table 3.6 contains the cross-classification cross-classification of hospital discharges by DCG discretion class and the three-category Roos index of discretion. Overall there was agreement in the two classification systems for over 41 percent of hospital discharges, and there was complete disagreement for only less than 1 percent of hospital discharges. This rate of complete high/low disagreement with the DCG discretion classifications was lowest by far among the three alternative

Table 3.5: Cross-Classification of DCG Discretion Score Classifications with Anderson Physican Discretion Classifications

	Anderson Physician Discretion Class				
DCG Discretion Class	Low Discretion	Moderate Discretion	High Discretion	Total Discharges	Marginal Distribution
Low Discretion	493.23 (18.61%)[1]	234.85 (8.86%)	1.34 (0.05%)	729.42	27.52%
Moderate Discretion	446.56 (16.85%)	379.00 (14.30%)	53.24 (2.01%)	878.80	33.16%
High Discretion	358.34 (13.52%)	669.77 (25.27%)	13.95 (0.53%)	1042.0	39.32%
Total Discharges	1298.14	1283.62	68.52	2650.3	
Marginal Distribution	48.98%	48.43%	2.59%		2650.3 100%

Percent of Agreement: 33.44%
Percent of Complete Disagreement: 13.57%

[1] Cell percentages of total discharges are reported in parentheses.

Table 3.6: Cross-Classification of DCG Discretion Score Classifications with Aggregated Roos Index of Discretion Classifications

	Roos Index of Discretion Class				
DCG Discretion Class	Low Variation	Moderate Variation	High Variation	Total Discharges (thousands)	Marginal Distribution
Low Discretion	211.30 (7.97%)[1]	508.27 (19.18%)	9.86 (0.37%)	729.42	27.52%
Moderate Discretion	70.45 (2.66%)	633.20 (23.89%)	175.16 (6.61%)	878.80	33.16%
High Discretion	10.46 (0.39%)	777.41 (29.33%)	254.19 (9.59%)	1042.0	39.32%
Total Discharges (thousands)	292.2	1918.88	439.21	2650.3	
Marginal Distribution	11.03%	72.40%	16.57%		100%

Percent of Agreement: 41.45%
Percent of Complete Disagreement: 0.76%

[1] Cell percentages of total discharges are reported in parentheses

discretion classification systems. Similar to our findings for the Anderson patient variation and physician discretion classifications, least agreement was found in classifying high discretion diagnoses. Only about 25 percent of hospital discharges classified as high discretion under the DCG discretion ratings were similarly classified under the aggregated Roos index of discretion.

In summary, the cross-classifications of all study hospital discharges by DCG discretion class and each of three alternative discretion classification systems revealed only modest levels of agreement between classifications. Overall agreement rates ranged between 33 percent (Anderson physician discretion) and 43 percent (Anderson patient variation). Complete disagreement high/low rates were relatively low, ranging from 0.8 percent (Roos index of discretion) to 13 percent (Anderson physician discretion). There were consistent patterns in relative agreement rates with individual DCG high, moderate, and low discretion classes. Agreement rates were uniformly highest for hospital discharges classified as low discretion under the DCG model. Consistently, the lowest level of agreement was found for hospital discharges classified as high discretion under the DCG model.

While it is comforting that the alternative discretion classification systems agree for a substantial fraction of hospital discharges, there clearly are differences that may not have marginal impacts. Furthermore, the finding that agreement rates were lowest for discharges classified as high discretion under the DCG model is of particular importance, since it is high discretion hospitalizations which were reclassified from higher risk classification and higher capitation rates to lower ones in the DCG model of Ellis and Ash (1995).

Overall, the level of disagreement among the classification systems are of sufficient magnitude that a comparison of the empirical performance of the alternative classification systems in distinguishing sources of HMO-FFS and geographic variations in hospital utilization is meaningful.

3.4.2 HMO-FFS Differences in Distribution of Discretionary Hospitalizations: A Comparison of Alternative Discretion Classification Systems

In the previous chapter, the distribution of hospital discharges and days of care among DCG discretion classes were compared for Medicare risk HMO and FFS as a means of testing the hypothesis that lower hospital use rates in Medicare risk HMOs relative to Medicare FFS use rates are attributable to HMO reductions in discretionary hospital use. If lower Medicare HMO use rates are largely the result of reductions in discretionary hospital use (as defined by discretion classification) then HMOs should have disproportionately fewer high discretion hospitalizations and days of care and disproportionately more low discretion hospitalizations than the Medicare FFS sector. The empirical findings reported earlier in Table 2.7 did not provide evidence of any greater reductions in discretionary hospital use in Medicare risk HMOs than nondiscretionary hospital use as defined under the DCG discretion ratings. This section contains empirical findings for the same HMO-FFS comparisons under the three alternative discretion classification systems.

Table 3.7 contains the distribution of hospital discharges and days of care among discretion classes for the Anderson patient variation classification system by Medicare risk HMO and FFS sectors for each of the study states. Similar to our findings reported earlier for the DCG discretion ratings in Table 2.7 these data provide little empirical support for the main study hypothesis. While Medicare risk HMOs had smaller percentages of high discretion hospitalizations in California and Florida, the higher percentage of high discretion hospital discharges and days of care for Medicare risk HMO enrollees in Massachusetts and New York contradict the main study hypothesis. With the exception of Massachusetts, there were a higher percentage of low discretion hospital discharges among risk HMO hospitalizations. The percentage of low discretion hospital days of care was higher for risk HMO hospitalizations than FFS hospitalizations in Massachusetts, however. While the HMO-FFS differences

Table 3.7: Distribution of Hospital Discharges and Days of Care among Anderson Patient Variation Classes by HMO/FFS and State

	State of Residence									
	California		**Florida**		**New York**		**Massachusetts**		**Total**	
Percent by Patient Variation Class +	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**
Percent of Discharges										
% Low Discretion	45.8	47.0	47.1	47.2	42.4	43.1	43.8	43.0	43.1	46.6
% Moderate Discretion	37.9	38.3	37.1	37.9	40.7	39.5	40.6	40.3	38.9	38.4
% High Discretion	16.3	14.7	15.8	14.9	15.6	17.4	15.6	16.7	16.2	15.0
Percent of Hospital Days										
% Low Discretion	48.3	50.6	49.2	51.8	46.7	49.0	48.0	48.8	17.4	50.8
% Moderate Discretion	35.5	33.0	34.9	33.7	37.5	35.2	38.0	37.1	36.4	33.4
% High Discretion	17.9	15.0	15.9	14.5	15.8	15.8	14.0	14.1	16.3	14.8
Total discharges (in thousands)	880	91	590	69	718	10	279	12	2468	182
Hospital discharge rate (/1,000)	359	189	304	266	330	140	362	272	338	221

+ The original Anderson patient variation classes of high, moderate, and low admission probability are inversely recoded as low, moderate, and high discretion, respectively.

in Table 3.7 are statistically significant given the large number of study hospital discharges, the practical HMO-FFS differences are only marginal.

Table 3.8 contains the empirical findings for the Anderson physician discretion classification system. Again the empirical findings are not supportive of the main study hypothesis. In fact, under the Anderson physician discretion definition of high discretion, a smaller percentage of high discretion risk HMO hospital discharges are only found in Florida and California, and those differences are only marginal. Similarly, the HMO-FFS differences in the distribution of days of care among discretion level classes are only marginal.

Table 3.9 contains the empirical findings for the aggregated Roos index of discretion. The largest HMO-FFS differences that were consistent with the main study hypothesis were those in New York. The findings for both hospital discharges and days of care were contradictory to expectations in Massachusetts. In general, however, the empirical findings are very similar to those of all of the other classification systems. There are only modest HMO-FFS and state level differences in the composition of hospital discharges and days of care among the classes of the Roos index of discretion, despite the fact that variation classes were determined empirically.

3.5 DISCUSSION

In this chapter three alternative discretion classification systems in the health services research literature were introduced as alternatives to the DCG discretion classifications. Comparisons of the discretion classification systems showed that for a majority of hospital discharges in the study data there was moderate (i.e., high versus moderate or low versus moderate) or complete (i.e., high versus low) disagreement between the DCG discretion classification and each alternative system. Nevertheless, the empirical findings regarding HMO-FFS differences discretionary/non discretionary hospital composition were quite similar to those reported earlier in Chapter 2 for the DCG model discretion

Table 3.8: Distribution of Hospital Discharges and Days of Care among Anderson Physician Discretion Classes by HMO/FFS and State

	State of Residence									
	California		**Florida**		**New York**		**Massachusetts**		**Total**	
Percent by Discretion Class	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**	**FFS**	**HMO**
Percent of Discharges										
% Low Discretion	50.0	50.1	50.1	48.4	46.9	48.4	46.5	46.5	49.0	49.4
% Moderate Discretion	47.4	47.7	46.7	48.4	50.1	49.2	51.2	50.6	48.4	48.3
% High Discretion	2.6	2.2	2.4	2.3	3.0	3.2	2.3	2.9	2.6	2.3
Percent of Hospital Days										
% Low Discretion	54.5	55.8	55.4	55.2	51.9	54.6	50.6	51.5	53.3	55.2
% Moderate Discretion	43.6	42.3	42.9	43.2	46.2	43.5	47.7	46.7	44.9	43.0
% High Discretion	2.0	1.9	1.7	1.6	1.9	1.9	1.7	1.8	1.8	1.8
Total discharges (in thousands)	880	91	590	69	718	10	279	12	2468	182
Hospital discharge rate (/1,000)	359	189	304	266	330	140	362	272	338	221

Table 3.9: Distribution of Hospital Discharges and Days of Care among Roos Index of Discretion Classes by HMO/FFS and State

	State of Residence									
	California		Florida		New York		Massachusetts		Total	
Percent by Level of Variation Class	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO	FFS	HMO
Percent of Discharges										
% Low Variation	11.0	12.3	11.7	13.1	11.1	12.2	11.4	11.1	11.3	12.5
% Moderate Variation	73.7	71.7	71.6	70.0	70.9	70.2	70.2	68.3	71.6	70.8
% High Variation	16.3	16.0	16.7	16.9	18.0	17.6	18.4	20.6	17.1	16.7
Percent of Hospital Days										
% Low Variation	11.7	13.2	13.1	15.5	13.2	15.0	13.7	13.4	12.8	14.2
% Moderate Variation	77.4	76.2	74.7	73.0	74.7	74.8	72.8	72.5	75.0	75.6
% High Variation	10.9	10.6	12.2	11.5	12.1	10.2	13.5	14.1	12.2	11.2
Total discharges (in thousands)	880	91	590	69	718	10	279	12	2468	182
Hospital discharge rate (/1,000)	359	189	304	266	330	140	362	272	338	221

+ The original Roos index of discretion categories have been aggregated as follows: high variation= Roos high variation I & II, low discretion= Roos low discretion, moderate discretion= all other discharges.

classes. Only very modest HMO-FFS differences in the distribution of hospital discharges and days of care among discretion classes were found regardless of which classification system was employed. The uniformity of the empirical findings indicate that at this level of data aggregation at least, the alternative discretion classification systems of Anderson, et al. (1989) and Roos, et al. (1988) do not appear to be any better than the DCG discretion classes advanced in the work of Ellis and Ash (1995).

The empirical findings reported in this chapter do not resolve the basic question of the validity of the discretion ratings employed in the DCG model of Ellis and Ash (1995). However, the analyses do indicate that whatever flaws may exist in the DCG model discretion ratings, similar flaws may be present in other discretion classification systems in the health services research literature. The next chapter takes a different approach to the issue by focusing on geographic variations in hospital utilization rates and the discretionary composition of hospitalizations.

GEOGRAPHIC VARIATIONS IN HOSPITAL USE RATES AND DISCRETIONARY HOSPITAL USE:

CHAPTER 4

4.1 INTRODUCTION

In the previous chapter we compared the distribution of hospital discharges and hospital days of care among discretion categories for Medicare FFS beneficiaries and Medicare risk HMO enrollees with several alternative measures of discretion in the literature. The data showed only very modest differences in the diagnostic composition of hospitalizations by broad high, moderate, and low discretion class between risk HMO and FFS sectors and among the four study states. While the basis for the relative invariance in the broad diagnostic composition of hospitalizations is uncertain, a plausible argument can be made that the empirical findings may be the result of the dominance of geographic practice style variations over HMO-FFS differences. If individual physicians are strongly influenced by the medical practice styles of their peers practicing in the same market areas, then HMO physicians may also be influenced by the prevailing medical practice styles of local FFS physicians. Furthermore, with the possible exception of staff model HMOs, the line distinguishing HMO from FFS physicians may be very blurry as it is common for many physicians to serve both FFS and HMO patients. Hence, the modest HMO-FFS differences in the discretionary composition of hospitalizations may be attributable at least in part to the common influence of medical practice styles that affect providers from both FFS and HMO delivery systems.

This chapter is focused on the analysis of geographic variations in hospital utilization rates with respect to the discretionary composition of hospitalizations. Exploratory multivariate analysis of geographic variations in hospital use rates are reported in Chapter 5. Given the modest HMO-FFS differences found in the previous chapter and concerns about the reliability of the primary payer code used to assign hospital discharges to Medicare risk HMO versus FFS sectors for several study states,

no distinctions are made between HMO and FFS beneficiaries in the analyses of this chapter. The empirical performance of the DCG discretion classifications will be compared with the same three alternative discretion classification systems introduced in the previous chapter.

4.2 GEOGRAPHIC VARIATIONS IN HOSPITAL UTILIZATION

Marked geographic variations in medical care utilization and costs have been documented in numerous studies conducted to date (Paul-Shaheen, et al. 1987; Folland and Stano 1990). While most studies have analyzed the utilization of the non-Medicare population, significant variations in both utilization and costs have also been found for the Medicare population (Gornick 1982, Deacon, et al. 1979, McClure and Shaller 1984, Rossiter and Adamache 1988, Roos, et al. 1988). While some researchers have disputed its significance (Moore 1985; Folland and Stano 1989, 1990), the most widely accepted explanation offered in past research is that observed variations are largely the result of professional discretion concerning treatment alternatives (Wennberg, et al. 1982; Wennberg 1985,1987; Wennberg, et al. 1987, Wennberg, et al. 1989).

The studies of Wennberg et. al (1987) and Wennberg et. al (1989) are two efforts that have addressed the contribution of hospitalizations for "high variation" medical conditions in explaining market differences in aggregate hospital utilization rates. Wennberg et. al (1987) found that the 50 percent higher overall hospital utilization rate in Boston, Massachusetts versus New Haven, Connecticut was largely due to higher discharge rates for "high variation/low consensus" medical conditions, such as pneumonia, gastroenteritis, and chronic obstructive lung disease. Discharge rates for "low variation/consensus" medical conditions, such as myocardial infarction and stroke, were nearly equal in the two areas. The later work of Wennberg et. al (1989) extended this research by subsequent analysis of 30 day hospital-associated mortality rates for Boston and New Haven. Their finding of no statistically significant difference in Medicare population-based death rates suggested that

the lower New Haven admission rate for high variation medical conditions was not associated with adverse mortality outcomes.

While the empirical findings of Wennberg et. al (1987) provides important evidence in support of the notion that large contribution of "high variation" medical conditions explains the higher hospital admission rates in Boston versus New Haven, it is not necessarily true that application of the same methodology to hospital use rates for a sample of metropolitan areas larger than two would yield consistent findings, let alone strong findings, among all pairs of geographic markets.

The work of Chassin et. al (1987) represents a more clinical approach for investigating whether geographic variations in physician use rates of surgical procedures might be explained by practice style differences. Detailed clinical data were collected for samples of Medicare patients receiving one of three surgical procedures (coronary angiography, carotid endarterectomy, and upper GI tract endoscopy) in three geographic sites distinguished by high, low, and moderate Medicare per capita use rates for the each of the three procedures. Employing clinical criteria developed by an expert panel of physicians, rates of appropriateness were calculated for the sites with varied rates for the surgical procedures. While some statistically significant differences among sites were found that were suggestive of high use sites having more inappropriate procedures than low use sites, the differences were always much too small to potentially explain the large differences in population use rates among the sites.

Subsequent research by Leape et. al (1990) expanded upon the earlier research by Chassin et. al (1987) to investigate the correlation between the percent of admissions judged to be inappropriate and the admission rate for the same three procedures for a sample of 23 counties adjacent to each other. A small but significant positive correlation was only found for coronary angiography. It was concluded that very little of the variation in the use rate for the three procedures could be attributed to

inappropriate use. Rather, high and low use counties tended to have roughly equal fractions of inappropriate admissions.

While Griffith et. al (1985) did not specifically investigate the association between physician discretion or appropriateness and hospital use rates, they did test whether diagnostic composition of hospital admissions was related statistically to local population hospital use rates for 60 market areas in the state of Michigan. Hospital admission mix was defined in several ways, including: body system of the primary diagnosis, selected surgical procedure groups, length-of-stay groups, and a grouping of common versus rare primary diagnoses on the basis of annual statewide admission frequencies. High use communities were found to have higher fractions of their total admissions with a respiratory disease diagnosis, with very common (versus rare) diagnoses, and with moderate (versus long or short) length-of-stay admissions. High use communities also tended to have smaller fractions of admissions with diagnoses of carcinoma and diseases of the nervous system, with rare diagnoses, and with long lengths-of-stay. While these compositional differences were found to be statistically significant, the differences were really quite modest, and could not lend much support to the proposition that high use communities had high use because of their mix of admissions. In fact, the principal finding of the study was that higher use communities tended to admit proportionately more patients in all of the patient mix categories than did lower use communities. In other words, the composition of hospital admissions in higher and lower total use areas was found to be quite similar.

Porell and Tompkins (1993) conducted an exploratory analysis of the contribution of discretionary admissions as reflected in the DCG discretionary scores to regional variations in hospital admission rates among a national sample of 176 geographic markets comprised of one or more MSA or non-MSA counties. The study which employed data for a five percent random sample of Medicare beneficiaries found some modest support for the premise that discretionary admissions, as defined in the DCG model of Ellis and Ash (1995), partially account for higher overall hospital admission rates in

some geographic markets. The quartile containing regions with the highest hospital admission rates on average had percentages of high discretion admissions that were about eight percent higher than those of regions in the quartile with the lowest hospital admission rates. Multivariate regression analysis revealed that a little over 12 percent of the variance in regional hospital discharge rates is accounted for by regional differences in the percentage of hospital episodes classified as high, moderate, or low discretion under the DCG discretionary ratings.

Other than adjusting hospital use rates for age-sex differences in population composition, Porell and Tompkins (1993) did not control for regional health status differences (e.g., through age-sex adjusted population mortality rates) in relating overall hospital use rates with the distribution of hospitalizations among DCG discretion score categories. Since higher hospital utilization for low discretion medical conditions would be expected in regions with populations of poorer health status, their empirical findings are likely to understate the contribution of high discretion hospitalizations to observed geographic variations in overall hospital use rates.

The geographic analyses reported in this chapter extend the work of Porell and Tompkins (1993) by the use of age-sex adjusted death rates as a population health status indicator. Furthermore, the current study employs a much larger sample of carefully defined geographic units that should better reflect potential systematic influences of medical practice styles on hospital use rates than the geographic units comprised of one or more county jurisdictional units in the Porell and Tompkins (1993) study.

4.3 DATA AND METHODOLOGY

Since the data sources have already been discussed in Chapter 2 of this report, discussion in this section is limited to the methodology employed in the delineation of geographic units and population characteristics of the units employed in the geographic analyses.

4.3.1 Study Geographic Units

While the 5-digit zip code is the smallest geographic identifier that is common to both hospital discharge and Medicare enrollment data files, relatively few zip codes contained a sufficient Medicare population to serve as geographic units for the study. Geographic market areas for the study were constructed by clustering 5-digit zip codes through a "hospital choice area" algorithm developed by Porell et. al (1991). The underlying premise for the clustering algorithm is that zip codes where patients tend to use the same hospitals are more likely to exhibit utilization patterns influenced by common medical practice styles. Under this premise, the clustering of zip codes with very similar patterns of usage among hospitals should produce larger geographic units that are relatively homogenous in terms of within-unit variance of zip code level utilization or expenditure rates. Porell et. al (1991) found hospital choice areas produced by this algorithm to have greater internal homogeneity (i.e., smaller relative within-unit variance of Medicare reimbursements per capita) than five alternative geographic units comprised of multiple 5-digit zip codes.

There were some basic requirements for the study geographic units that guided the process of their delineation. First, it was important that the individual geographic units be spatially cohesive and all units together be spatially exhaustive. Second, a minimum threshold population of about 5,000 Medicare beneficiaries was sought for each unit since this population size is sufficient for estimating Medicare hospital use rates with standard errors that are less than 5 percent of the sample mean. Third, given that there are many fewer Medicare risk HMO enrollees than FFS beneficiaries generally, we sought to maximize the number of geographic study units with at least 1,000 Medicare risk HMO enrollees. Finally, we sought to maximize the number of geographic units to increase the study sample size for analysis of geographic variations. Given the absence of contiguity constraints in the clustering algorithm and the arbitrary nature of any stopping criterion, there was no assurance that a simple

application of the algorithm would produce geographic units satisfying these requirements. Accordingly, the study units were delineated through an interactive process that was strongly guided by the assignments produced by the clustering algorithm. Details of the process are summarized in Appendix A in this report.

The clustering algorithm was applied with patient origin counts of zip code-to-hospital discharges for all patients who were 65 years of age or older in a zip code. Out-of-state hospital discharges were not counted, and no distinction was made between HMO and FFS patients in the patient origin counts. The clustering algorithm was applied to all spatial zip codes with 2 or more discharges in the states of Massachusetts and Florida. Because of the much larger number of 5-digit zip codes in California and New York, these states were first divided into north and south regions and the algorithm was applied separately for each of the regions. The interactive process of geographic unit delineation led to the definition of a total of 767 geographic units in total for the four study states. This general sample was further reduced to 761 units for all geographic analyses by restricting these analyses to geographic units with 900 or more person-years of aged Medicare beneficiaries with Part A eligibility.

Table 4.1 contains descriptive information about the populations of these 761 geographic units for each of the study states. Examination of the table shows that in 1992 the mean aged Medicare population size for all study geographic units was 10,761 person-years. The population size distribution of geographic units, as reflected in the mean and standard deviation of Medicare aged population person-years, was very comparable among study states. The mean and standard deviation of Medicare risk HMO market penetration among study geographic units suggest a skewed geographic distribution of Medicare HMO risk enrollment in each of the study states, with many geographic units having few or no Medicare risk HMO enrollees.

Table 4.1: Population Statistics for Study Geographic Units with 900 or More Medicare Aged Beneficiaries in 1992 by State

	California	Florida	Massachusetts	New York	Total
Number of Geographic Units	281	194	67	219	761
Aged Medicare population *					
Mean	10528	11314	12022	10184	10761
Std. Deviation	6959	6600	9982	9460	7972
Minimum	1348	1059	1744	1414	1059
Maximum	51767	39733	52963	57563	57563
Risk HMO enrollees as % of total population					
Mean					
Std. Deviation	15.7%	10.8%	5.7%	2.8%	9.9%
Minimum	15.2%	12.2%	7.1%	5.2%	12.8%
Maximum	0.0%	0.0%	0.0%	0.0%	0.0%
	59.4%	44.0%	28.7%	27.9%	59.4%
Geographic Units with 900+ Medicare risk HMO enrollees	150	92	24	35	301
Risk HMO population					
Mean	3237	2788	1538	1612	2776
Std. Deviation	2214	1509	622	625	1893
Minimum	938	926	916	914	914
Maximum	13517	7737	3694	3725	31517
Geographic Units with 900+ Medicare FFS beneficiaries	281	194	67	219	761
FFS Medicare population					
Mean	8749	9963	11366	9833	9601
Std. Deviation	6155	5878	9572	9091	7415
Minimum	1252	1016	1735	1411	1016
Maximum	49729	39546	50689	54316	54316

* Population counts are measured in person-years of Part A eligibility.

The bottom portion of Table 4.1 presents HMO and FFS population size data for geographic units meeting the minimum 900 person-year population threshold when the aged Medicare population is stratified into separate risk HMO and FFS populations. In light of the skewed geographic distribution of a much smaller risk HMO population, there is a substantial reduction in the number of study geographic units with sufficient risk HMO populations for a separate analysis of geographic variations in hospital use. The reduction in geographic units was particularly large for the state of New York where only 35 of 219 study geographic units had Medicare risk HMO populations exceeding the 900 person-year threshold.

Table 4.2 contains general information about the distribution of age-sex adjusted hospital discharge rates and hospital days of care rates among geographic units in each of the study states. As expected, mean hospital use rates were lower for Medicare risk HMO enrollee populations relative to FFS populations. At the same time, however, the standard deviations of the hospital use rates were larger for risk HMO enrollees than those of the FFS Medicare population. While some of the greater geographic variability of risk HMO hospital use rates should be attributable, at least in part, to sampling variation associated with smaller population sizes, a closer examination of the risk HMO hospital use rate distributions raised serious questions about the reliability of the data fields used to assign hospital discharges to risk HMO versus FFS populations.

The hospital use rates at the 10th percentile of the geographic unit distribution for risk HMO populations reported in Table 4.2 show that there were a fairly large number of geographic units with extremely low risk HMO hospital use rates in all study states with the possible exception of Florida. For example, in California where the estimated age-sex adjusted hospital discharge rate was 331 discharges per 1,000 aged Medicare beneficiaries in 1992, there were 15 geographic units with risk HMO discharge rates as low or lower than 86 discharges per 1,000 person-years. There were four geographic units in New York with Medicare risk HMO hospital discharge rates less than 44 discharges

Table 4.2: Distributions of Hospital Discharge and Hospital Day Rates Among Study Geographic Units by State and by HMO/FFS

State/ HMO-FFS		Hospital discharge rate (/1,000)				Hospital days rate (/1,000)			
	Geographic Units	Mean	Std. dev.	10th%	90th%	Mean	Std. dev.	10th%	90th%
California	281	342	79	263	440	2591	1183	1691	3520
FFS	281	377	101	273	524	2902	1302	1817	4126
HMO	150	199	140	86	303	1266	1340	438	1992
Florida	194	317	50	257	394	2446	530	1899	3134
FFS	194	322	62	261	404	2551	603	1938	3356
HMO	92	265	50	212	316	1645	357	1253	1966
Massachusetts	67	347	46	257	409	2591	491	2211	3471
FFS	67	356	62	289	409	2980	521	2341	3533
HMO	24	316	178	52	489	2337	1371	372	3830
New York	219	331	53	271	399	3832	985	2843	5161
FFS	219	335	50	273	411	3890	1026	2845	5274
HMO	35	125	71	43	229	1319	781	444	2330
All States	761	333	65	263	412	2938	1106	1899	4127
FFS	761	349	50	263	454	3103	1147	1972	4351
HMO	301	220	127	82	328	1473	1112	505	2441

per 1,000 beneficiaries. In the combined sample of 301 geographic units from all study states with risk HMO populations exceeding 900 person-years, there were 31 geographic units with adjusted hospital discharge rates less than 83 discharges per 1,000 population.

Hospital use and mortality rates were examined more carefully for both risk HMO and FFS populations in these "low outlier" geographic units to assess whether the very low risk HMO hospital use rates might be plausibly explained by factors such as random variation associated with small population size, or population health status differences reflected in low mortality rates. Table 4.3 contains unweighted sample means for hospital use, mortality rate, and population size for four groups of geographic units: (1) risk HMO enrollees in any one of the 31 low risk HMO use geographic units; (2) FFS beneficiaries residing in any one of the same 31 geographic units; (4) risk HMO enrollees in all other geographic units; and (4) FFS beneficiaries in all other geographic units. Examination of the risk HMO population means in Table 4.3 shows that mean risk HMO population size in the lowest use rate geographic units was actually larger than the mean of all other 270 geographic units meeting risk HMO population threshold requirements. Hence, the low HMO hospital use rates do not appear to be concentrated among geographic units with the smallest risk HMO populations. In fact, the geographic unit with the second lowest risk HMO annual hospital discharge rate of 18 hospital discharges per 1,000 beneficiaries had a risk HMO population of 10,951 person-years. While it is interesting that the mean FFS beneficiary population in the same geographic units was nearly double that of the residual 730 geographic units with FFS populations exceeding 900 person-years, these data did not suggest that the low risk HMO hospital use was attributable to random variation due to small population size.

Adjusted hospital use rates for FFS beneficiaries in the 31 low outlier geographic units were then compared with the FFS hospital use rates of other geographic units. While the mean FFS hospital use rate of 374 discharges per 1,000 beneficiaries was much higher than the mean rate of 52 hospital discharges per 1,000 population for Medicare HMO enrollees in the same geographic units, these FFS

Table 4.3: A Comparison of Risk HMO and FFS Hospital Use and Mortality Rates in the Lowest HMO Use Rate Geographic Units Versus All Other Geographic Units

Subgroup of Geographic Units		Mean Population (person-years)	Mean Adjusted Annual Hospital Discharge Rate (/1,000)	Mean Adjusted Annual Days of Care Rate (/1,000)	Mean Adjusted Annual Mortality Rate
Risk HMO Enrollees	Low Outlier Geographic Units (n=31)	2914	52	442	2.59%
	All Other Units (n=730)	2760	239	1592	3.50%
FFS Beneficiaries	Low Outlier Geographic Units (n=31)	17451	374	3406	4.89%
	All Other Units (n=730)	9267	348	3091	4.74%

hospital use rates were not unusually low or high in comparison to the mean FFS hospital use rate of 348 discharges per 1,000 beneficiaries in the other residual study geographic units. Finally, a comparison of mean risk HMO and FFS adjusted mortality rates in these low outlier geographic units did not reveal differences that could practically account for the large disparity between HMO-FFS hospital use rates in these geographic units. The mean adjusted annual mortality rate for HMO enrollees in these 31 geographic units of 2.59 percent was much smaller than the mean mortality rate of 4.89 percent for FFS beneficiaries in the same geographic units. However, the mean mortality rate of 3.50 percent for risk HMO enrollees in all other geographic units with higher hospital use rates was also low in comparison to FFS mortality rates.

Given these findings, the most plausible explanation for the extremely low Medicare risk HMO hospital use rates in a number of geographic units is misclassification of hospitalizations. As discussed earlier in Chapter 2, hospital discharges were assigned to risk HMO versus FFS Medicare populations on the basis of expected primary payer fields in the state discharge files. The misclassification of HMO hospital discharges to FFS populations is much more likely than the misclassification of FFS hospital discharges to risk HMO enrollees given the dominance of Medicare as the primary payer for hospitalizations among the elderly. Since aged Medicare FFS beneficiaries roughly outnumber aged Medicare risk HMO enrollees in the study states by a factor of 9 to 1, the misclassification of risk HMO enrollee hospital discharges should have a much greater effect on the estimated hospital use rates of smaller risk HMO populations than larger FFS populations.

As noted earlier, vendors of the hospital discharge data files were questioned about known systematic coding problems with the expected primary payer field. While the vendors acknowledged that Medicare HMO hospital discharges may be undercounted among the elderly due to the dominance of Medicare as the primary payer of hospitalizations, the vendors were unaware of systematic reporting problems for specific hospitals except in Massachusetts where serious HMO undercounting was found

for a subset of geographic units served by hospitals with known reporting problems for the payer field. Given the absence of any objective criteria for reliably identifying specific geographic units with misclassified hospital discharges to be discarded from the study samples, geographic analysis of HMO and FFS-specific hospital use rates will not be conducted due to serious concerns over the reliability of the assignment of hospital discharges to risk HMO versus FFS populations. Geographic analyses of hospital use rates are restricted to the aggregate utilization of combined FFS and risk HMO Medicare beneficiary populations.

4.3.2 Study Research Hypotheses

The main research hypotheses to be addressed through the geographic analysis reported in this chapter may be stated as follows:

H1: Hospital use rates for medical conditions rated as high (low) discretion in terms of physicians' decisions to hospitalize will exhibit larger (smaller) geographic variations in hospital discharge and days of care rates.

H2: Geographic markets with higher (lower) overall rates of hospital use by the aged Medicare population should have disproportionately more (fewer) hospitalizations for medical conditions rated as high discretion and disproportionately fewer (more) hospitalization rated as low discretion.

The basis for the first study hypothesis is the "professional uncertainty" thesis advanced by Wennberg et. al (1982). Under this thesis it is maintained that there is physician discretion toward hospitalization and/or the performance of certain surgical procedures because the relative clinical outcomes of treatment alternatives may not be well established. Since physicians may disagree among themselves concerning the most appropriate course of care for a given medical condition, the decision to hospitalize may be highly discretionary. Given this lack of consensus about what is the appropriate treatment, local physician practice styles regarding the use of hospital treatment may be influenced not

only by general physician attitudes, but also by factors such as the availability of hospital beds (Wennberg et. al 1987).

Under this thesis it can be deduced that the level of geographic variability in hospital discharge rates for medical conditions should vary inversely with the degree of professional consensus about the appropriateness of acute hospitalization as a course of treatment. Smaller geographic variations have been found for conditions such as myocardial infarction and inguinal hernia repair, where current medical standards generally dictate inpatient hospital treatment, and greater variations have been found for surgical procedures where there has been controversy about their value, such as hysterectomy, prostatectomy, and tonsillectomy, (e.g., Wennberg and Gittelsohn 1982, McPherson et. al 1982, Roos et. al 1988). The latter hospitalizations have been described as "high variation" in the literature. Common usage of the term "high variation" in the geographic variations literature has evolved to imply more than just the degree of observed variations in utilization, but also the lack of consensus about the effectiveness of treatment alternatives.

While the motivation for the variability hypothesis H1 is found in the small area geographic variations literature, there is an important difference in our research approach. The DCG discretion classes were defined on an a priori basis through physician ratings of diagnoses. Geographic variability is used as a means for evaluating the validity of the physician ratings regarding level of physician discretion.

Even if greater geographic variability in hospital use rates are found for hospitalizations classified as higher discretion, it cannot necessarily be concluded that the regional variations in overall hospital use rates are largely attributable to discretionary hospital use. It is also necessary to show that high variation discharges account for a disproportionately larger share of hospitalizations in geographic markets with higher overall hospital use rates. This is the motivation for the second study hypothesis, H2.

4.3.3 Methodology

Discretion Classifications

All four discretion classification systems to be compared were specified in terms of three-category ordinal variables to facilitate comparisons of empirical performance. DCG discretion categories of high, moderate, and low discretion were specified in terms of ranges of the total discretion scores by principal diagnosis of the hospitalization as discussed in Chapter 2. The original three-category physician discretion classifications of Anderson et. al (1989) by principal diagnosis were employed without modification. As discussed earlier in Chapter 3, the patient variation classifications of admission probabilities were inversely recoded into high, moderate, and low discretion categories that should be positively associated with the other discretion measures. Finally, the aggregated three-category ordinal index of discretion of Roos et. al (1988), developed in Chapter 3, was also employed in the geographic analyses.

Geographic Variability and Discretion Level

The study hypothesis H1 was addressed by comparing percentage weighted coefficients of variation (WCVs)[7] for the distribution of high discretion and low discretion age-sex adjusted hospital discharge rates among geographic units. Geographic units served as cases which were weighted by the ratio of person-years of 1992 Part A Medicare eligibility in the geographic unit to the mean person-years for all geographic units in the state or pooled sample from all states. WCVs were only computed for high and low (and not moderate) discretion/variation subgroups of hospitalizations under the premise that the greatest differences in variability are expected for these extreme subgroups.

The percentage coefficient of variation (CV) is defined as the (standard deviation/sample mean) x 100. The percentage WCV differs from a simple CV only in that cases are not equally weighted.

While WCVs naturally adjust for much of observed differences underlying prevalence rates for various subgroups of hospital discharges, Diehr et. al (1993) have shown that they still exhibit some modest negative correlation with prevalence rates. Diehr et. al (1993) advocate the use of CVs derived from person-level analysis of variance models to measure relative variability for subgroups of hospitalizations since they have been shown to be invariant with respect to prevalence rates. However, they also note that such estimates are based on assumptions that may not be true (e.g., variance in the number of admissions per person is constant among geographic areas). In any case, analysis of variance CVs could not be estimated in this study since multiple discharges for individual patients could not be distinguished in the study hospital discharge data. The theoretical shortcomings of WCVs as measures of geographic variability may not matter much, however, as Diehr et. al (1993) note that with relatively larger populations all of the typical variance measures employed in the small area variations literature essentially yield similar rank ordering of variability among subgroups of hospital discharges.

Overall Hospital Use Rates and the Discretionary Mix of Hospitalizations

The second main study hypothesis H2 was addressed by correlating overall hospital discharge and days of care rates with variables measuring the percentage of high discretion, moderate discretion, and low discretion hospital discharges. Both simple Pearson correlations and partial correlations, in which age-sex adjusted annual population mortality rates were controlled, were estimated. Empirical support for the study hypothesis can be found in positive (negative) correlations between overall age-sex adjusted hospital use rates and the percentage of high (low) discretion hospitalizations. If annual mortality rates serve as a useful marker of population health status differences which are not reflected in the age-sex population composition, then stronger partial correlations are expected once mortality rate differences are controlled.

4.4 EMPIRICAL RESULTS

4.4.1 Weighted Coefficients of Variation

Table 4.4 contains the empirical findings for percentage WCVs for hospital discharges classified either as high or low discretion under the four discretion classification systems. WCVs were computed separately for geographic units in each study state and for the combined pool of geographic units in all four states.[8] Under the study hypothesis H1, hospital discharge rates for high discretion hospitalizations should exhibit greater geographic variability than low discretion hospitalizations. This should be reflected in larger WCVs.

The WCVs for hospital discharge rates in the high and low discretion categories for the DCG model discretion classes in Table 4.4 do not provide consistent empirical support for the study hypothesis H2. WCVs for high discretion hospitalizations were only a little larger than those of low discretion hospitalizations in three of the study states. In contrast to expectations, the WCV was larger for DCG low discretion hospitalizations than for high discretion hospitalizations in California. Finally, there was virtually no difference between the high and low DCG model discretion group WCVs for the pooled sample of geographic units for all study states.

The WCVs for the three alternative discretion classifications lend modest, but consistent empirical evidence in support of the study hypothesis H1. WCVs were uniformly higher for high discretion hospital discharge subgroups relative to low discretion subgroups. Comparisons among alternative classification systems reveals two important patterns in the findings. The WCVs for high and low variation subgroups of the Roos index of discretion suggest that hospital discharge rates for the Roos high variation subgroup exhibit greater geographic variability and discharge rates for the Roos low variation subgroup and exhibit lesser geographic variability than their counterpart subgroups under

---

Note that since the mean hospital discharge rates vary among the study states, the WCVs for the pooled sample of geographic units will not necessarily reflect the weighted average of state WCVs.

Table 4.4: Percentage Weighted Coefficients of Variation for High and Low Discretion Hospitalizations Under Alternative Classification Systems

State	High/Low Discretion Class	DCG Discretion Scores	Anderson Physician Variation	Anderson Physician Discretion	Roos Index of Discretion
California N=281	Low	24.7%[1]	20.9%	21.2%	16.8%
	High	22.8%	25.2%	34.6%	27.3%
Florida N=194	Low	17.6%	17.8%	18.1%	16.5%
	High	19.4%	21.2&	30.5%	24.8%
Massachusetts N=67	Low	10.9%	11.7%	10.9%	9.8%
	High	12.7%	14.6%	28.9%	17.4%
New York N=229	Low	12.6%	13.0%	13.6%	12.2%
	High	16.5%	17.2%	37.9%	20.5%
All States N=771	Low	20.3%	18.2%	18.3%	16.1%
	High	20.2%	22.0%	37.1%	24.7%

[1] Percentage WCVs are defined as (std. dev/mean)*100 for cases weighted by their relative person-years of Part A eligibility.

both the DCG model and the Anderson patient variation classifications systems. Since the Roos index of discretion was developed empirically on the basis of observed geographic variations for individual modified DRG groups of hospitalizations, this finding is not surprising.

A more puzzling pattern in the WCV findings are the very large WCVs found for the high discretion subgroup of hospitalizations under the Anderson physician discretion classifications. The WCVs were consistently much greater than those for the Roos high variation subgroup. Given the very small fractions of high discretion hospitalizations under the Anderson physician discretion classes, these higher WCVs may be attributable in part to bias in the WCV measurement of variability. However, since Diehr et. al (1993) showed that WCVs had only a slight negative correlation with prevalence rates of hospitalizations, this is unlikely to account fully for the higher level of geographic variability in discharge rates for the Anderson high physician discretion subgroup of hospitalizations. The relatively few diagnoses that were distinguished as high discretion in the physician discretion classification of Anderson et. al (1989) could simply comprise a "small tail" subgroup of hospitalizations with hospital discharge rates exhibiting the highest levels of geographic variability.

Overall, the empirical findings relating geographic variability to level of discretion were generally supportive of the hypothesis that high discretion hospitalizations should exhibit greater geographic variability in hospital discharge rates than low discretion hospitalizations. Since only age-sex demographic adjustments were made to hospital discharge rates, the geographic variability should also reflect uncontrolled population health status differences. However, if low discretion hospital use is more sensitive to population health status differences than is high discretion hospital use, comparisons of WCVs for high and low discretion discharge subgroups are more likely to understate than overstate differences in geographic variability that are not associated with population health status differences.

The comparison of WCVs for alternative discretion classification systems also indicate that the Anderson discretion classifications, and the Roos index of discretion in particular, are superior to the

DCG discretion classes in distinguishing among hospitalizations exhibiting higher and lower levels of geographic variability. Comparisons of WCVs do not allow for conclusions about there being an "excess of high variability hospitalizations" in geographic areas with higher overall hospital use rates WCVs. For such a conclusion to be made, discharge rates for high variability hospitalization subgroups must be systematically higher (lower) in geographic areas with higher (lower) overall hospital use. This is an empirical question for which the correlation analysis should provide some insight about.

4.4.2 Correlations between Overall Hospital Use Rates and the Discretionary Composition of Hospitalizations

Under one of the main study hypotheses, H2, high discretion hospitalizations should account for a disproportionately larger share of hospitalizations in geographic areas with high overall hospital use rates. Empirical support for this study hypothesis can be found in positive and negative correlations between the overall hospital utilization rate and the percentage of high discretion and low discretion hospital discharges, respectively. Empirical findings for both simple and partial correlations in which population mortality rates differences are controlled are reported in this section.

DCG Model Discretion Classes

Table 4.5 contains empirical findings for the DCG model discretion classification system. Table 4.5 shows that for the pooled sample and for two of the individual states, small but statistically significant positive simple and partial correlations were found between overall hospital use rates and the percent of high discretion discharges. Similarly, modest but significant negative correlations were found for the percent of low discretion hospital discharges for several states. While controlling for differences in mortality rates had little impact on the correlations for shares of high discretion hospital

Table 4.5: Pearson Simple and Partial Correlations Between Total Hospital Utilization Rates and the Percent of Hospital Discharges in DCG Discretion Classes by State

	Hospital Discharge Rate (/1,000)		Hospital Days Rate (/1,000)	
State / DCG Discretion Class	Pearson Correlation	Partial Correlation / Adjusted Mortality Rate	Pearson Correlation	Partial Correlation/ Adjusted Mortality Rate
California (N=281)[1]				
Low Discretion	-0.06	-0.03	-0.06	-0.04
Moderate Discretion	-0.19**	-0.16**	-0.20**	-0.19**
High Discretion	0.22**	0.17**	0.22**	0.18**
Florida (N=194)				
Low Discretion	-0.19**	-0.36**	-0.17**	-0.30**
Moderate Discretion	0.10**	0.29**	0.17**	0.38**
High Discretion	0.06	0.03	-0.02	-0.09
Massachusetts (N=67)				
Low Discretion	-0.38**	-0.39**	-0.20**	-0.19**
Moderate Discretion	0.31**	0.33**	0.02	0.07
High Discretion	0.06	0.04	0.15**	0.18**
New York (N=219)				
Low Discretion	-0.45**	-0.45**	-0.32**	-0.31**
Moderate Discretion	0.21**	0.37**	0.30**	0.43**
High Discretion	0.17**	0.06	0.00	-0.10**
All States (N=761)				
Low Discretion	-0.11**	-0.14**	-0.08	-0.11**
Moderate Discretion	-0.07	0.05	=0.15**	-0.05
High Discretion	0.17**	0.09**	0.21**	0.14**

[1] Cases are weighted by relative Part A Medicare population in person-years.

** $P < 0.05$

discharges, in most instances, the partial correlations between overall hospital use rates and the percent of low discretion admissions were more negative than the simple correlations. This suggests that low discretion hospital use may be more sensitive to health status differences reflected in mortality rates than higher discretion use. However, at best, the overall empirical findings lend only modest empirical support for the study hypothesis H2 under the DCG discretion classification system.

Anderson Patient Variation Classification

Table 4.6 contains the correlation results for the Anderson patient variation classification system. Overall, the correlations for the patient variation ratings followed a pattern similar to the DCG discretion class findings. The correlations for the Anderson patient variation ratings were a little stronger than those for the DCG discretion classes. In addition, the general pattern of the direction and size of the correlation among the states was a little more consistent than for the DCG model ratings. Again, much stronger (negative) correlations were found between overall hospital use rates and the percent of low discretion hospitalizations than for high discretion hospitalizations. This emerging pattern in the correlation results suggests that there may be a greater distinction between low discretion hospitalizations and other hospitalizations than between high discretion hospitalizations and others.

Anderson Physician Discretion Classification

Table 4.7 contains the empirical findings for the physician discretion classification of Anderson et. al (1989). While it did not hold true for Florida, the most consistent pattern in the empirical findings is a negative correlation between percentage of low discretion hospital discharges and overall hospital use rates. Contrary to expectations, significant negative correlations were found for the percentage of high discretion hospitalizations in California. The rather marginal empirical findings for the Anderson physician discretion classification are notable in light of the extremely high level of

Table 4.6: Pearson Simple and Partial Correlations Between Total Hospital Utilization Rates and the Percent of Hospital Discharges in Anderson Patient Variation Classes by State

	Hospital Discharge Rate (/1,000)		Hospital Day Rate (/1,000)	
State / Anderson Patient Variation Class	Pearson Correlation	Partial Correlation / Adjusted Mortality Rate	Pearson Correlation	Partial Correlation/ Adjusted Mortality Rate
California (N=281)[1]				
Low Discretion	-0.24**	-0.23**	-0.31**	-0.30**
Moderate Discretion	0.15**	0.13**	0.12**	-0.09
High Discretion	0.16**	0.13**	0.26**	0.17**
Florida (N=194)				
Low Discretion	-0.20**	-0.17**	-0.21**	-0.19**
Moderate Discretion	0.16**	0.13**	0.16**	0.17**
High Discretion	0.08	0.08	0.06	0.05
Massachusetts (N=67)				
Low Discretion	-0.19**	-0.18**	-0.11**	-0.09
Moderate Discretion	0.23**	0.23**	0.07	0.06
High Discretion	-0.02	-0.03	0.06	0.05
New York (N=219)				
Low Discretion	-0.41**	-0.18**	-0.38**	-0.19**
Moderate Discretion	0.42**	0.47**	0.32**	0.35**
High Discretion	0.05	0.09	0.14**	0.17**
All States (N=761)				
Low Discretion	-0.28**	-0.21**	-0.60**	-0.58**
Moderate Discretion	0.27**	0.20**	0.51**	0.47**
High Discretion	0.10**	0.08	0.30**	-0.30**

[1] Cases are weighted by relative Part A Medicare population in person-years.

** $P < 0.05$

Table 4.7: Pearson Simple and Partial Correlations Between Total Hospital Utilization Rates and the Percent of Hospital Discharges in Anderson Physician Discretion Classes by State

	Hospital Discharge Rate (/1,000)		Hospital Day Rate (/1,000)	
State / Anderson Physician Discretion Class	Pearson Correlation	Partial Correlation / Adjusted Mortality Rate	Pearson Correlation	Partial Correlation/ Adjusted Mortality Rate
California (N=281)[1]				
Low Discretion	-0.18**	-0.10**	-0.18**	-0.15**
Moderate Discretion	0.18**	0.13**	0.18**	0.16**
High Discretion	-0.12**	-0.10**	-0.09	-0.00
Florida (N=194)				
Low Discretion	-0.15**	-0.03	-0.09	0.05
Moderate Discretion	0.18**	0.05	0.18**	-0.00
High Discretion	-0.07	-0.07	-0.07	-0.00
Massachusetts (N=67)				
Low Discretion	-0.40**	-0.39**	-0.16**	-0.13**
Moderate Discretion	0.38**	0.37**	0.15**	0.12**
High Discretion	0.11**	0.10**	0.06	0.05
New York (N=219)				
Low Discretion	-0.32**	-0.32**	-0.23**	-0.22**
Moderate Discretion	0.38**	0.25**	0.24**	0.19**
High Discretion	-0.01	0.13**	-0.04	0.05
All States (N=761)				
Low Discretion	-0.25**	-0.15**	-0.49**	-0.44**
Moderate Discretion	0.27**	0.16**	0.46**	0.39**
High Discretion	-0.04	-0.02	0.17**	0.22**

[1] Cases are weighted by relative Part A Medicare population in person-years.

** $P < 0.05$

geographic variability that was found for discharge rates of high discretion hospitalizations under this discretion classification (see Table 4.3). Apparently the geographic areas with higher overall hospital use rates do not systematically have higher use rates for this discretion class subgroup.

Roos Index of Discretion

Table 4.8 contains empirical findings for the Roos index of discretion. Comparing the Roos classification findings with the others already presented in Table 4.5 through Table 4.7 reveals that the correlations for the Roos classification are uniformly stronger than those of the other discretion classifications. The empirical findings also exhibit the greatest consistency among states. With only a few exceptions, correlations between the percentage of high variation hospital discharges and overall adjusted hospital discharge rates were uniformly positive and statistically significant. All correlations between the percentage of low variation discharges and hospital discharge and days of care rates were negative and significant.

4.4.3 Multiple Regression Model Results

A quantitative assessment of how well the Roos classification performs empirically, relative to the other discretion classifications is not easy to see from the WCV and correlation analyses. Conventional R-square measures of model fit for linear regression models can provide a scale for quantifying the relative performance of the discretion classifications. This section reports findings from a set of exploratory multiple regression analyses performed to assess the fraction of variance in age-sex adjusted overall hospital discharge and days of care rates accounted for by differences in the discretionary composition of hospital discharges and age-sex adjusted mortality rates.

Table 4.8: Pearson Simple and Partial Correlations Between Total Hospital Utilization Rates and the Percent of Hospital Discharges in Roos Index of Discretion Classes by State

	Hospital Discharge Rate (/1,000)		Hospital Day Rate (/1,000)	
State / Roos Index of Discretion Class	Pearson Correlation	Partial Correlation / Adjusted Mortality Rate	Pearson Correlation	Partial Correlation/ Adjusted Mortality Rate
California (N=281)[1]				
Low Variation	-0.60**	-0.60**	-0.54**	-0.53**
Moderate Variation	0.10**	0.27**	0.19**	0.27**
High Variation	0.35**	0.27**	0.22**	0.15**
Florida (N=194)				
Low Variation	-0.38**	-0.28**	-0.39**	-0.29**
Moderate Variation	0.03	0.02	0.14**	0.17**
High Variation	0.19**	0.14**	0.10**	0.02
Massachusetts (N=67)				
Low Variation	-0.60**	-0.68**	-0.45**	-0.44**
Moderate Variation	0.10**	0.12**	-0.19**	-0.17**
High Variation	0.35**	0.45**	0.53**	0.51**
New York (N=219)				
Low Variation	-0.60**	-0.62**	-0.57**	-0.61**
Moderate Variation	0.01	0.13	0.30**	0.42**
High Variation	0.40**	0.34**	0.11	0.03
All States (N=761)				
Low Variation	-0.54**	-0.53**	-0.44**	-0.42**
Moderate Variation	0.04	0.13**	-0.06	0.02**
High Variation	0.31**	0.22**	0.34**	0.26**

[1] Cases are weighted by relative Part A Medicare population in person-years.

** $P < 0.05$

Ordinary least squares regression analysis, with cases weighted by relative person-years of Part A eligibility, was employed to fit a set of multiple regression models in each state. Three independent variables were specified in these models:

- the percentage of high discretion hospital discharges,
- the percentage of low discretion hospital discharges,
- the age-sex adjusted mortality rate for the aged Medicare population.

The percentage of moderate discretion discharges was omitted to avoid complete multicollinearity with the model constant. The adjusted mortality rate was specified in each model as a marker variable to capture population health status differences not reflected in the age-sex demographic composition. Higher adjusted hospital use rates are expected in geographic areas with higher mortality rates. In accord with the study hypothesis H2 and the omission of percent of moderate discretion discharges, the percentage of high discretion discharges should be positively associated with higher hospital use rates. The percentage of low discretion hospitalizations should exhibit a negative association.

Before discussing the empirical findings, a word of caution is exercised concerning the model specification. Since high and low discretion hospitalizations along with moderate discretion hospital discharges comprise the overall total of hospital discharges, the percentage high and low discretion compositional variables are not truly exogenous. Hence, the individual estimated regression coefficients should be biased due to simultaneous equations bias. However, the general pattern of empirical findings and R-square measures of model fit will still provide useful information for comparing among the alternative discretion classifications.

Tables 4.9 and 4.10 contain all estimated regression coefficients and R-square model fit measures for models with total adjusted hospital discharges and days of care rates as dependent variables, respectively. In general, the estimated coefficients are not uniformly consistent with a priori

Table 4.9: Empirical Results for the Regression of Total Adjusted Hospital Discharge Rate on Discretion Class Composition and Population Mortality Rate Variables

	Discretion Classification System			
	DCG Discretion Scores	Anderson Physician Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
California				
% High Discretion	931 (3.5)*	65 (0.2)	-407 (2.0)*	15 (0.1)
% Low Discretion	475 (1.9)	-686 (3.1)*	-1139 (2.0)*	-3086 (11.1)*
Mortality Rate	51 (5.4)*	54 (5.9)*	53 (5.6)*	45 (5.6)*
Constant	-390 (2.5)*	395 (2.8)*	326 (2.8)*	476 (7.9)*
Adjusted R-square	0.146	0.155	0.129	0.427
Florida				
% High Discretion	-417 (3.0)*	-44 (0.2)	-118 (0.9)	-25 (0.2)
% Low Discretion	-1072 (6.2)*	-317 (2.2)*	-629 (1.2)	-931 (3.5)*
Mortality Rate	61 (16.1)*	57 (14.0)*	57 (13.6)*	54 (13.2)*
Constant	485 (5.6)*	211 (2.3)*	129 (1.6)	184 (3.4)*
Adjusted R-square	0.588	0.521	0.510	0.545
Massachusetts				
% High Discretion	125 (0.8)	-66 (0.3)	-319 (0.8)	258 (2.0)*
% Low Discretion	-572 (3.6)*	-331 (2.2)*	-529 (4.1)*	-906 (5.9)*
Mortality Rate	7 (1.4)	7 (1.1)	5 (0.9)	4 (0.9)
Constant	215 (2.4)*	266 (2.9)*	373 (5.3)*	184 (4.7)*
Adjusted R-square	0.259	0.071	0.199	0.525
New York				
% High Discretion	-682 (4.9)*	-475 (2.5)*	285 (1.0)	380 (2.7)*
% Low Discretion	-1670 (9.2)*	-1049 (8.5)*	-655 (4.6)*	-1890 (10.3)*
Mortality Rate	43 (8.5)*	41 (8.3)*	39 (6.8)*	36 (7.9)*
Constant	843 (9.2)*	658 (8.6)*	441 (5.7)*	299 (7.8)*
Adjusted R-square	0.402	0.380	0.253	0.503
All States				
% High Discretion	-356 (2.3)*	843 (4.4)*	1253 (3.6)*	-857 (6.3)*
% Low Discretion	-325 (1.9)	193 (1.7)	604 (5.5)*	-2731 (13.6)*
Mortality Rate	46 (8.8)*	43 (8.4)*	52 (9.9)*	40 (8.5)*
Constant	324 (3.4)*	-115 (1.4)	-258 (3.8)*	580 (13.4)*
Adjusted R-square	0.089	0.106	0.122	0.252

Table 4.10: Empirical Results for the Regression of Ttal Adjusted Hospital Days Rate on Discretion Class Composition and Population Mortality Rate Variables

	Discretion Classification System			
	DCG Discretion Scores	Anderson Physician Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
California				
% High Discretion	11286 (3.7)*	5943 (1.9)	-3568 (0.5)	-4856 (1.9)
% Low Discretion	5666 (2.0)*	-8069 (3.3)*	-6122 (2.7)*	-34355 (10.3)*
Mortality Rate	414 (3.9)*	450 (4.4)*	454 (4.2)*	427 (4.5)*
Constant	-5370 (2.9)*	3201 (2.0)*	3582 (2.6)*	5213 (7.3)*
Adjusted R-square	0.106	0.158	0.084	0.337
Florida				
% High Discretion	-6304 (4.6)*	-44 (0.2)	-118 (0.9)	-25 (0.2)
% Low Discretion	-10930 (6.3)*	-317 (2.2)*	-629 (1.2)	-931 (3.5)*
Mortality Rate	557 (14.8)*	57 (14.0)*	57 (13.6)*	54 (13.2)*
Constant	5310 (6.2)*	211 (2.3)*	129 (1.6)	184 (3.4)*
Adjusted R-square	0.550	0.521	0.510	0.545
Massachusetts				
% High Discretion	710 (0.1)	-77 (0.1)	582 (0.1)	11480 (3.1)*
% Low Discretion	-5473 (1.2)	-2375 (0.6)	-3335 (0.9)	-8886 (1.9)
Mortality Rate	370 (2.4)*	367 (2.4)*	353 (2.3)*	270 (2.1)*
Constant	2533 (1.0)	2404 (1.0)	2957 (1.5)	709 (0.6)
Adjusted R-square	0.080	0.052	0.059	0.332
New York				
% High Discretion	17908 (6.3)*	-1485 (0.4)	1013 (0.2)	9341 (3.3)*
% Low Discretion	-28819 (8.0)*	-16398 (6.4)*	-9249 (3.2)*	43614 (11.9)*
Mortality Rate	717 (6.9)*	611 (6.0)*	547 (4.7)*	664 (7.3)*
Constant	15710 (8.4)*	8298 (5.3)*	5700 (3.6)*	7413 (9.7)*
Adjusted R-square	0.307	0.259	0.135	0.454
All States				
% High Discretion	5055 (2.7)*	-908 (0.5)	8787 (2.3)*	5216 (3.2)*
% Low Discretion	-1602 (0.8)	-18875 (16.7)*	-14408 (11.9)*	-26630 (10.5)*
Mortality Rate	732 (11.7)*	566 (11.1)*	597 (10.5)*	612 (10.7)*
Constant	-1954 (1.8)	9007 (11.3)*	7047 (9.4)*	2267 (4.4)*
Adjusted R-square	0.190	0.450	0.336	0.325

* P < 0.05

expectations. With only a few exceptions, mortality rates and hospital use rates are shown to have positive and statistically significant associations. Also, the percentage of low discretion hospitalizations was generally found to have negative and significant associations with overall hospital discharge and days of care rates. However, the estimated coefficients for the percentage of high discretion discharges variable were inconsistent. Relatively few of these coefficients were positive and significant, and a number of them were actually negative and significant.

A comparison of the R-square measures among discretion classification systems on Tables 4.9 and 4.10 generally shows rather modest explanatory power for the models specified with the DCG or Anderson discretion classification variables. In most cases the R-squares of the models specified with the Roos classification variables had much higher R-squares than the other three models. A comparison among models suggests that consistent large (in absolute value) negative coefficients for the percentage of low discretion discharges accounts for much of the additional explained variance for the models with the Roos index of discretion specified. This general pattern suggests that the stronger empirical performance of the Roos classifications are largely the result of its superior classification of low discretion/low variation hospitalizations than its classification of high discretion/variation hospitalizations. However, the relatively modest statistical fits for the models do not provide strong empirical support for the premise that higher overall hospital use rates in certain geographic areas are attributable in large part to excessive high discretion hospitalizations, as defined under the four alternative classification systems.

4.5 DISCUSSION

In this chapter geographic analyses were conducted to test the main study hypotheses regarding the relative geographic variability of high versus low discretion hospitalizations, and the degree to which higher overall rates of hospital use are attributable to excessive high discretion hospitalizations.

Analysis of WCVs for subgroups of high and low discretion hospitalizations for the various discretion classification systems provide some empirical support for the premise that the geographic variability of hospital discharge rates for medical conditions should vary inversely with the level of discretion associated with treatment through hospitalization. Correlation and regression analyses findings suggest higher use rate geographic areas tend to have a greater fraction of hospitalizations classified as high discretion/variation, but that the higher overall hospital use rates are not in large part the result of excessive high discretion hospitalizations.

A comparison of the relative empirical performance of the DCG discretion classification system with those of Anderson et. al (1989) and Roos et. al (1988) revealed only modest differences among the DCG and Anderson discretion classifications derived from a priori physician ratings of diagnoses. The Roos index of discretion derived empirically on the basis of consistent patterns of geographic variability exhibited superior empirical performance relative to the other classification systems. However, the performance was not at a such a high level that would permit strong conclusions of the sort drawn by Wennberg et. al (1987) about high discretion hospitalizations. For a sample of two metropolitan areas, Wennberg et. al (1987) concluded that hospitalizations for "high variation" medical conditions accounted for virtually all of the fifty percent higher overall hospital discharge rate in Boston relative to New Haven. An expanded sample of geographic areas in this study did not reveal such distinctive impacts of high discretion hospital use on geographic variations in overall hospital use rates.

In fact, the empirical analyses suggest that it is not hospitalizations classified as "high discretion" that are strongly associated with higher overall hospital use rates. Rather, higher overall hospital use rates appear to be much more strongly associated with excessive hospitalizations for all conditions <u>not</u> classified as "low discretion." The large negative correlations between high overall hospital use rates and the percentage of low discretion hospital discharges suggest that high and low overall hospital use rate areas will have much more similar absolute low discretion hospital use rates

than for all other hospitalizations. The correlation analyses for all four discretion classifications indicated that the fraction of low discretion hospitalizations served as a much more effective marker for distinguishing among geographic areas with higher or lower overall hospital use rates than the fraction of high discretion hospitalizations.

The empirical results of this chapter suggest that a relatively small share of hospitalizations are likely to be truly nondiscretionary in the sense that hospital use rates are relatively invariant once population health status differences are controlled. Such findings have implications for refinements for prior use capitation payment models. If the share of discretionary hospitalizations that are vulnerable to medical practice style differences is not relatively small, reducing the influence of medical practice styles on prior use risk classifications by modifying the risk classifications of discretionary hospitalizations will not be an effective means of establishing a risk classification system that is both relatively invariant to medical practice styles and predictive of relative risks for future medical expenses.

The correlation and regression findings regarding the association between overall hospital use rates and the percentage of hospital discharges by discretion class have implications about the similarity or differences in absolute hospital use rates for high and low discretion hospitalizations among geographic areas with differing overall hospital use rates. The next chapter contains further empirical analysis of geographic variations in absolute hospital use rates for high and low discretion hospitalizations and their association with market area supply factors.

DISCRETIONARY HOSPITAL USE AND SUPPLY FACTORS

CHAPTER 5

5.1 INTRODUCTION

The empirical analyses reported in the previous chapter indicated that low discretion hospitalizations exhibited smaller geographic variations than other hospitalizations, and that areas with higher overall hospital use rates tended to have smaller fractions of low discretion hospitalizations. Under the "professional uncertainty hypothesis" of Wennberg, et al. (1982), greater physician discretion is largely the result of disagreement among physicians concerning the most appropriate course of care for a given medical condition. Given this lack of consensus about appropriate treatment, local physician practice styles regarding the use of hospital treatment may be influenced not only by general physician attitudes, but also by supply factors such as the availability of hospital beds (Wennberg, et al. 1987). Physicians practicing in geographic markets where more hospital beds are available may be more inclined to hospitalize patients for conditions for which there is greater discretion over the course of treatment than those in markets where the bed supply is tighter.

Aggregate hospital utilization rates have commonly been found to be more highly correlated with measures of physician and hospital bed supply than with population health status (e.g., Wilson and Tedeshi 1984; Wennberg and Gittelsohm 1982; McPherson et. al, 1981; Pasley et. al, 1987). However researchers have not explicitly tested some obvious hypotheses concerning the linkages between the professional uncertainty hypothesis, discretionary and nondiscretionary hospital use, and physician and hospital bed supply. More specifically, it is reasonable to hypothesize that discretionary hospital use rates will be strongly associated with market supply factors than nondiscretionary hospital use rates.

In this chapter, the above hypothesis and others will be tested through some simple multivariate regression models of factors associated with low and high discretion hospital discharge rates. Separate

models are estimated for each of the four alternative discretion classification systems analyzed in the previous chapters. The next section contains an overview of the hypotheses to be tested, the general methodology, data sources, and model specification. The empirical findings are reported in the fourth section. The last section contains a discussion of the findings and their policy implications.

5.2 HYPOTHESES AND MODEL DEVELOPMENT

5.2.1 Hypotheses

There are two explicit hypotheses related to the general study hypotheses stated in Chapter 1. Each hypothesis will be stated in general terms of the dichotomy of discretionary versus nondiscretionary hospitalization. The four discretionary classification systems studied in previous chapters are employed to operationalize these subgroups of hospitalizations. The first hypothesis may be stated as follows:

Ha: Adjusted hospital discharge rates for discretionary hospitalizations will be more highly associated with physician and hospital bed supply variables than low discretion hospitalizations.

The implicit premise underlying much of the empirical small area geographic variations literature is that physician and hospital bed supply influence physicians' decisions to hospitalize for medical conditions where there is greater discretion in choice of treatment alternatives. While there is disagreement about the underlying causes of commonly observed positive correlations between hospital use rates and supply measures of physician and hospital beds (e.g., see Folland and Stano 1990), it is clearly plausible that a physician's decision to opt for hospitalization over some alternative course of treatment for a patient is easier to make if there are hospital beds readily available.

While it could be posited that nondiscretionary hospital use should be invariant with respect to physician and hospital supply factors, this is not required for empirical support of the hypothesis, Ha above. Depending on the level of health status-driven population demand for inpatient hospital

treatment relative to supply, nondiscretionary hospital use may also be correlated with supply factors. Under Ha we are only hypothesizing that discretionary hospital use will be more sensitive than nondiscretionary hospital use to the influence of supply factors. No hypotheses are advanced about the relationship between nondiscretionary use rates and supply factors.

A second hypothesis to be tested concerns HMO-FFS differences in medical practice styles. If HMO physicians exhibit more conservative practice styles than their FFS counterparts, it is reasonable to posit that HMO discretionary hospital use rates will be less sensitive to the influence of supply factors than FFS discretionary hospital use rates. Furthermore, if certain hospitalizations are truly nondiscretionary, then HMO and FFS nondiscretionary hospital use rates should not vary once population health status differences are controlled.

Concerns over the reliability of expected primary payer codes on hospital discharge file records preclude a direct test of risk HMO and FFS medical practice style differences on discretionary and nondiscretionary hospital use. However, it is possible to distinguish among geographic areas where higher/lower fractions of the aged Medicare population are Medicare risk HMO enrollees. The aggregate hospital use rates of the combined risk HMO and FFS aged Medicare beneficiary populations should reflect the hospital use of both risk HMO and FFS populations. Accordingly, if there were no HMO-FFS medical practice style differences, and population health status differences among geographic areas are adequately controlled, differences in the percentage of Medicare risk HMO enrollees among geographic areas should have no association with hospital use rate differences. However, if risk HMOs are relatively successful in reducing discretionary hospital use due to more conservative medical practice styles, lower discretionary hospital use rates are expected in geographic areas with larger fractions of their population receiving care as risk HMO enrollees. Since nondiscretionary hospitalizations are posited to be less vulnerable to medical practice styles, no

differences are expected in nondiscretionary hospital use since population health status differences are adequately controlled.

It should be noted that it is also possible that there may be competitive effects of HMOs on FFS providers in geographic areas with high HMO market penetration. That is, the practice styles of FFS providers in areas with high HMO market penetration may be more conservative as well. If so, FFS beneficiaries in such areas may also exhibit lower rates of discretionary hospital use. Since only hospital use rates of the combined population of risk HMO and FFS beneficiaries are analyzed, such competitive effects cannot be distinguished in this analysis.

The formal study hypothesis may be stated as follows:

Hb: Adjusted hospital use rates for discretionary hospitalizations will be more negatively associated with the percentage of Medicare risk HMO enrollees in the population than with nondiscretionary hospitalizations.

Again it should be noted that empirical support for Hb does not require that absolute discretionary hospital use rates be lower in geographic areas where there are greater levels of Medicare risk HMO market penetration. It is well known that Medicare HMO risk contracting has been selective with respect to higher levels of AAPCC capitation rates (Adamache and Rossiter 1986; Porell and Wallack 1990). Accordingly, there may be a positive correlation between hospital use rates and percentage of Medicare risk HMO enrollees due to systematic geographic biases in selective Medicare risk HMO enrollment. If a positive correlation exists between nondiscretionary hospital use rates and Medicare risk HMO market penetration, a less positive or negative correlation is expected for discretionary hospital use rates because of (expected) more conservative HMO medical practice styles.

It might be posited that there should be differences in the sensitivity of discretionary and nondiscretionary hospital use rates to population health status differences. Alternatively it can be argued that population health status differences should only affect the prevalence of medical conditions in a population and that medical practice factors (independently) determine whether the level of

physician discretion associated with treatment options. Since both positions can be reasonably defended, no formal hypothesis is stated with respect to differences in the influence of population health status differences on discretionary versus nondiscretionary hospital use.

5.2.2 General Model Specification

The simplest general linear model to be estimated is stated below:

$$\mathbf{NONDISC}_i = \alpha_0 + \alpha_1\,\mathbf{HLTH}_i + \alpha_2\,\mathbf{HMO}_i + \alpha_3\,\mathbf{MARKET}_i + \epsilon_i\,, \qquad [1]$$

$$\mathbf{DISC}_i = \beta_0 + \beta_1\,\mathbf{HLTH}_i + \beta_2\,\mathbf{HMO}_i + \beta_3\,\mathbf{MARKET}_i + \mu_i\,. \qquad [2]$$

DISC and NONDISC are defined as age-sex adjusted hospital discharge rates for discretionary and nondiscretionary hospitalizations in each geographic area i. HLTH is defined generally as one or more measures of population health status (e.g., age-sex adjusted mortality rate) for the geographic area other than age-sex population composition. HMO is defined as the percentage of population in geographic area i who are Medicare risk HMO enrollees. MARKET is defined as one or more measures characterizing the level or composition of market supply factors (e.g., hospital beds per capita) for geographic area i. The terms ϵ_i and μ_i represent random disturbances.

A significant positive coefficient is expected for the coefficient of MARKET in the discretionary hospitalization equation given past research findings about positive associations between overall hospital use rates and physician and/or hospital bed supply variables (e.g., Wilson and Tedeshi 1984; Wennberg and Gittelsohm 1982; McPherson et. al, 1981; Pasley et. al, 1987). No particular association is theoretically expected between nondiscretionary hospital use rates and MARKET, once health status effects are controlled for variables specified in HLTH. Under the specific formal hypothesis Ha above, however, the key expectation is that the estimated coefficient for α_3 of MARKET

in the nondiscretionary hospital use equation will be smaller than β_3 of MARKET in the discretionary hospital use equation, or $(\beta_3 - \alpha_3) > 0$.[1]

Since theoretically HMOs should only be more successful than the FFS sector in reducing discretionary admissions, under the second formal hypothesis, Hb, the estimated coefficient for α_2 of HMO in the nondiscretionary hospital use equation should be greater than β_2 of HMO in the discretionary hospital use equation, or $(\beta_2 - \alpha_2) < 0$. However as discussed above, the absolute magnitudes of the estimated coefficients of the HMO market penetration variable are uncertain.

Finally, if the health status variables comprising HLTH are scored such that a higher value represents a lower level of population health status (e.g. age-sex adjusted mortality rates), a significant positive coefficient is expected for the estimated coefficient of α_1 in the nondiscretionary equation. As discussed above, it is unclear whether discretionary hospital use will vary with population health status. Hence the expected sign and significance for β_1 of HLTH in the nondiscretionary equation is uncertain.

5.2.3 Variable Specification

Discretionary and nondiscretionary hospital discharges were defined as the high discretion/variation and low discretion/variation categories of the DCG, Anderson, and Roos discretion classifications introduced in earlier chapters. With the specification of alternative discretion classifications their relative performance can be assessed in terms of consistency with the hypothesized relationships and overall model fit.

In the linear model specifications in equations [1] and [2] above, the estimated regression coefficients reflect the expected marginal impact on hospital discharge rates associated with a unit

---

For this assertion to be technically correct there must be equal prevalence rates for discretionary and nondiscretionary hospitalizations. As discussed later, discharge rates were standardized to remove the influences of different prevalence rates.

increase in the independent variables. If high and low discretion subgroups have different prevalence rates, the study hypotheses cannot be tested by comparing estimated regression coefficients. Accordingly, absolute age-sex adjusted annual hospital discharge rates for high and low discretion subgroups were standardized through z-scores based on the state (or pooled) sample means and standard deviations of their respective subgroups to remove effects of different prevalence rates.

Population Health Status

Age and gender differences in population composition are already controlled through direct age-sex adjustments to hospital discharge rates. Two additional variables were specified as indicators of population health status for the HLTH term in equations [1] and [2]. The age-sex adjusted population annual mortality rates (MORTALITY RATE) were specified as one of these indicators. Geographic areas with higher adjusted mortality rates are expected to have higher discharge rates of nondiscretionary hospitalizations. The percentage of the population with Medicaid buy-in status (PERCENT BUY-IN) was specified as a variable reflecting potential health status differences associated with income differences among populations. This variable was technically measured as the percentage of person-years of Part A eligibility with Medicaid buy-in status. Higher nondiscretionary hospital discharge rates are expected in geographic areas with larger shares of beneficiaries with Medicaid buy-in status.

Market Supply Factors

The Area Resource File (ARF) contains a wide range of variables for characterizing health delivery system markets and supply factors at the county level. The study geographic units were defined by clustering five-digit zip codes with similar hospital choice patterns. While most of the geographic areas are much smaller than counties, this in itself is not a problem since multiple

geographic units within a single county can obviously be assigned identical values for market attribute variables from the ARF. However, since many of the ARF variables commonly specified in empirical studies have relatively high levels of intercorrelation at county level units of observation, the intercorrelations among many ARF variables are even higher for the study population where there are multiple geographic units within counties.[2] Given that the study hypotheses advanced in this chapter are tested by comparing the magnitude of estimated regression coefficients, unstable parameter estimates associated with multicollinearity among market variables would make such comparisons very unreliable.

Since it was important that the discharge rate models [1] and [2] have stable regression coefficients, we chose to avoid potential multicollinearity problems by specifying the MARKET factors in discharge equations [1] and [2] employing 1992 data from the ARF. Specifically, two different model specifications were tested. In the first model, a single physician supply variable, PHYSICIANS/CAPITA, defined as the number of active nonfederal physicians per 100 persons of total county population in 1992, was specified for MARKET. In the second model specification, a single hospital bed supply variable, HOSPITAL BEDS/CAPITA, defined as the number of short term hospital beds per 100 persons of total county population in 1992, was specified. For geographic units with Medicare populations in more than one county, the variables were specified as weighted average of the county-level ARF variables, where the weights were the fractions of Part A aged Medicare population of the geographic unit in each county.

It would have been desirable to specify both hospital and physician supply variables as well as other market area variables from the ARF in a broader model. Some test runs for states where hospital

---

Potential multicollinearity problems of unstable coefficient estimates were compounded further by the pooled estimation of equations [1] and [2] which permitted statistical tests of parameter differences between equations. The pooled model specification is discussed below.

and physician supply variables were jointly tested exhibited patterns of unstable coefficients that were symptomatic of multicollinearity problems. At the risk of possible specification bias due to omission of important market area attribute variables, simple models with single market attribute variables were estimated for testing the study hypotheses.

5.2.4 Pooled Estimation of Discretionary and Nondiscretionary Models

In order to facilitate tests of the statistical significance of differences in the estimated parameters of the discretionary and nondiscretionary models [1] and [2], two observations were created for each geographic unit. For one observation, the dependent variable was specified as the high discretion hospital discharge rate. The low discretion hospital discharge rate was specified as the dependent variable for the other observation of a geographic unit. The two observations were pooled together into one larger analytical file used for parameter estimation.

A dummy variable (HIGH DISCRETION) was set equal to unity for observations for which the dependent variable was the discretionary hospitalization discharge rate, and zero otherwise. This dummy variable was specified as an independent variable, and was also interacted with all other independent variables in the pooled discharge rate model. Under this single pooled model version of equations [1] and [2], the constant and estimated coefficients for the main independent variables reflect estimates of the coefficients for the low discretion discharge model [2]. The estimated coefficient for the dummy variable HIGH DISCRETION and the remaining coefficients for all interaction variables constructed from the HIGH DISCRETION dummy variable are technically interpreted as differences between the estimated high and low discretion model coefficients, or (β_j - α_j) for the jth independent variable. A positive and statistically significant coefficient for any one of these interaction variables means that the estimated β coefficient for the discretionary hospital use equation [2] exceeds the

estimated α coefficient for the nondiscretionary hospital use rate equation, and the difference between coefficients is statistically significant at conventional levels of statistical significance.

5.3 EMPIRICAL RESULTS

Table 5.1 through 5.5 contain the empirical results for the 40 regression models estimated for individual study states and the pooled study population. The top portion of each table contains the estimated coefficients and t-statistics for the models specified with the physician supply market variable. The bottom portion contains results for the models specified with the hospital bed supply market variable. Before discussing the implications of estimated model coefficients toward the study hypotheses, some general patterns in the empirical results are noted.

The empirical regression results suggest that both high discretion and low discretion hospital discharge rates are associated with population health status. Except for Massachusetts where the estimated coefficients were not statistically significant, parameter estimates for the population health status variable MORTALITY RATE were uniformly positive and statistically significant. However with only a few exceptions, the estimated coefficient for the interaction variable MORTALITY RATE*HIGH DISCRETION was not significant. Together these findings suggest that both low and high discretion hospital discharge rates are similarly associated with population mortality rates. These results support the view that population health status differences affect the prevalence of medical conditions in a population regardless of the level of discretion associated with how a medical condition is treated by physicians.

The estimated coefficients for the other population health status surrogate variable, PERCENT-BUY IN, showed similar patterns to those for the mortality rate variables. The results suggest that low discretion hospital discharge rates are higher in geographic areas with greater concentrations of dually-eligible beneficiaries with Medicaid buy-in status. A number of

Table 5.1: Multiple Regression Results for Physician Supply and Hospital Supply Models: California

California (N=562)	DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
Physician Supply Model				
Constant	-285 (5.4)*	-320 (6.0)*	-313 (5.2)*	-360 (6.9)*
High Discretion	-52 (0.7)	46 (0.6)	319 (3.8)*	44 (0.6)
Mortality Rate	40 (3.4)*	44 (3.7)*	41 (3.0)*	59 (5.0)*
Mortality Rate*High Dis	5 (0.3)	-11 (0.6)	-57 (3.0)*	-14 (0.9)
Percent Risk HMO	120 (4.0)*	120 (4.0)*	119 (3.6)*	63 (2.2)*
Percent Risk HMO*High Dis	-21 (0.5)	-75 (1.8)	-119 (2.5)*	75 (1.8)
Percent Buy-in	270 (4.3)*	371 (5.9)*	393 (5.5)*	281 (4.5)*
Percent Buy-in*High Dis	159 (1.8)	32 (0.4)	20 (0.2)	153 (1.8)
Physicians/Capita	163 (3.4)*	167 (3.5)*	190 (3.5)*	159 (3.4)*
Physicians/Capita*High Dis	38 (0.6)	42 (0.6)	-170 (2.2)*	-61 (0.9)
Adjusted R^2	0.26	0.26	0.17	0.28
Hospital Bed Supply Model				
Constant	-286 (5.4)*	-315 (6.0)*	- 324 (5.4)*	-329 (6.3)*
High Discretion	-77 (1.0)	-15 (0.2)	269 (3.2)*	-19 (0.3)
Mortality Rate	37 (3.2)*	41 (3.6)*	39 (2.9)*	54 (4.7)*
Mortality Rate*High Dis	7 (0.4)	-7 (0.4)	-50 (2.7)*	-9 (0.5)
Percent Risk HMO	115 (3.9)*	117 (3.9)*	112 (3.3)*	65 (2.2)*
Percent Risk HMO*High Dis	-27 (0.6)	-89 (2.1)*	-125 (2.6)*	64 (1.5)
Percent Buy-in	251 (4.0)*	355 (5.7)*	366 (5.2)*	278 (4.5)*
Percent Buy-in*High Dis	143 (1.6)	-2 (0.0)	15 (0.1)	130 (1.5)
Hospital Beds/Capita	190 (3.4)*	178 (3.2)*	249 (3.9)*	99 (1.8)
Hospital Beds/Capita*High	111 (1.4)	214 (2.7)*	-59 (0.7)	98 (1.2)
Adjusted R^2	0.27	0.29	0.19	0.27

* $P < 0.05$

Table 5.2: Multiple Regression Results for Physician Supply and Hospital Supply Models: Florida

Florida (N=388)	DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
Physician Supply Model				
Constant	-468 (15.1)*	-439 (12.7)*	-432 (11.3)*	-335 (9.9)*
High Discretion	45 (1.0)	41 (0.8)	144 (2.7)*	10 (0.2)
Mortality Rate	95 (14.0)*	94 (12.4)*	89 (10.6)*	74 (10.0)*
Mortality Rate*High Dis	-7 (0.7)	-14 (1.3)	-39 (3.3)*	-9 (0.9)
Percent Risk HMO	37 (0.9)	2 (0.0)	1 (0.0)	137 (3.1)*
Percent Risk HMO* High Dis	-123 (2.2)*	-116 (1.8)	-81 (1.1)	-293 (4.7)*
Percent Buy-in	198 (3.3)*	148 (2.3)*	207 (2.8)*	38 (0.6)
Percent Buy-in*High Dis	-44 (0.5)	2 (0.0)	-240 (2.3)*	228 (2.5)*
Physicians/Capita	71 (1.0)	2 (0.0)	53 (0.6)	-84 (1.1)
Physicians/Capita*High Ds	-6 (0.1)	152 (1.4)	314 (2.6)*	176 (1.7)
Adjusted R^2	0.55	0.47	0.35	0.40
Hospital Bed Supply Model				
Constant	-487 (15.8)*	-459 (13.4)*	- 452 (11.7)*	-381 (11.3)*
High Discretion	56 (1.3)	54 (1.1)	177 (3.3)*	54 (1.2)
Mortality Rate	93 (13.8)*	93 (12.4)*	87 (10.3)*	74 (9.9)*
Mortality Rate* High Dis	-6 (0.6)	-16 (1.5)	-44 (3.7)*	-10 (0.9)
Percent Risk HMO	36 (0.9)	-7 (0.2)	-2 (0.0)	108 (2.6)*
Percent Risk HMO*High Dis	-119 (2.2)	-94 (1.6)	-35 (0.5)	-253 (4.3)*
Percent Buy-in	228 (4.3)*	152 (2.6)*	231 (3.5)*	14 (0.2)
Percent Buy-in*High Dis	-48 (0.6)	58 (0.7)	-125 (1.3)	287 (3.5)*
Hospital Beds/Capita	96 (2.5)*	64 (1.5)	89 (1.8)	93 (2.2)*
Hospital Beds/Capita*High	-36 (0.6)	45 (1.5)	77 (1.1)	-52 (0.8)
Adjusted R^2	0.56	0.47	0.34	0.41

* $P < 0.05$

Table 5.3: Multiple Regression Results for Physician Supply and Hospital Supply Models: New York

New York (N=438)	DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
Physician Supply Model				
Constant	-346 (7.1)*	-343 (6.8)*	-315 (5.7)*	-273 (5.5)*
High Discretion	37 (0.5)	84 (1.2)	312 (3.9)*	1 (0.0)
Mortality Rate	68 (6.9)*	68 (6.7)*	59 (5.3)*	64 (6.4)*
Mortality Rate*High Dis	-9 (0.7)	-27 (1.9)	-61 (3.9)*	-14 (0.9)
Percent Risk HMO	-54 (0.6)	-57 (0.6)	-39 (0.3)	-24 (0.3)
Percent Risk HMO*High Dis	-130 (0.9)	-120 (0.9)	-253 (1.6)	-284 (2.1)*
Percent Buy-in	481 (4.9)*	540 (5.3)*	663 (5.9)*	46 (0.5)
Percent Buy-in*High Dis	386 (2.8)	466 (3.2)*	-822 (5.2)*	-671 (4.7)*
Physicians/Capita	-7 (0.3)	2 (0.0)	-30 (1.2)	-113 (4.8)*
Physicians/Capita*High Ds	-67 (2.1)*	152 (1.4)	158 (4.3)*	69 (2.1)*
Adjusted R^2	0.36	0.35	0.21	0.32
Hospital Bed Supply Model				
Constant	-314 (6.3)*	-333 (6.5)*	- 321 (5.5)*	-291 (5.7)*
High Discretion	-17 (0.2)	22 (0.3)	395 (4.8)*	21 (0.3)
Mortality Rate	65 (6.7)*	70 (7.1)*	62 (5.6)*	72 (7.3)*
Mortality Rate* High Dis	-1 (0.0)	-23 (1.7)	-78 (5.0)*	-20 (1.4)
Percent Risk HMO	-76 (0.8)	-89 (0.9)	-51 (0.4)	-74 (0.7)
Percent Risk HMO*High Dis	-133 (0.9)	-58 (1.6)	-218 (1.4)	-259 (1.9)
Percent Buy-in	533 (5.2)*	583 (5.5)*	671 (5.6)*	84 (0.8)
Percent Buy-in*High Dis	344 (2.4)*	346 (2.3)*	-789 (4.6)*	662 (4.4)*
Hospital Beds/Capita	-53 (1.6)	-80 (2.4)*	-31 (0.8)	-124 (3.8)*
Hospital Beds/Capita*High	-16 (0.3)	149 (3.2)*	96 (1.8)	64 (1.4)
Adjusted R^2	0.36	0.36	0.17	0.31

* $P < 0.05$

Table 5.4: Multiple Regression Results for Physician Supply and Hospital Supply Models: Massachusetts

Massachusetts (N=134)	DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
Physician Supply Model				
Constant	-128 (1.5)	-121 (1.3)	-108 (1.0)	-147 (1.9)
High Discretion	13 (0.1)	15 (0.1)	-16 (0.1)	1 (0.0)
Mortality Rate	22 (1.1)	17 (0.8)	9 (0.4)	9 (0.5)
Mortality Rate*High Dis	-8 (0.3)	-12 (0.4)	6 (0.2)	-2 (0.1)
Percent Risk HMO	12 (0.1)	83 (0.7)	29 (0.2)	-121 (1.2)
Percent Risk HMO*High Dis	65 (0.4)	-62 (0.4)	-71 (0.4)	196 (1.4)
Percent Buy-in	89 (0.5)	137 (0.7)	240 (1.0)	123 (0.7)
Percent Buy-in*High Dis	301 (1.1)	309 (1.1)	-89 (0.3)	271 (1.1)
Physicians /Capita	143 (3.9)*	159 (4.1)*	175 (3.9)*	27 (0.8)
Physicians/Capita*High Ds	40 (0.8)	36 (0.7)	-35 (0.6)	99 (2.1)*
Adjusted R^2	0.30	0.29	0.17	0.33
Hospital Bed Supply Model				
Constant	-65 (0.7)	-51 (0.5)	-45 (0.4)	-102 (1.3)
High Discretion	-7 (0.1)	-36 (0.3)	-34 (0.2)	-6 (0.1)
Mortality Rate	10 (0.5)	5 (0.2)	-3 (0.1)	4 (0.2)
Mortality Rate* High Dis	-8 (0.3)	-10 (0.3)	9 (0.3)	-6 (0.3)
Percent Risk HMO	59 (0.5)	135 (1.1)	94 (0.7)	-127 (1.2)
Percent Risk HMO*High Dis	95 (0.6)	-21 (0.1)	-82 (0.4)	251 (1.7)
Percent Buy-in	179 (0.8)	238 (1.0)	280 (1.0)	299 (1.5)
Percent Buy-in*High Dis	146 (0.5)	14 (0.1)	-120 (0.3)	91 (0.3)
Hospital Beds/Capita	75 (1.2)	82 (1.3)	134 (1.8)	-81 (1.5)
Hospital Beds/Capita*High	129 (1.5)	211 (2.3)*	-13 (0.1)	198 (2.6)*
Adjusted R^2	0.16	0.20	0.04[1]	0.30

[1] F-statistic for overall regression model P=0.11

* $P < 0.05$

Table 5.5: Multiple Regression Results for Physician Supply and Hospital Supply Models: All States

All States (N=1522)	DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion
Variable	Coefficient (t)	Coefficient (t)	Coefficient (t)	Coefficient (t)
Physician Supply Model				
Constant	-364 (15.4)*	-329 (13.7)*	-331 (12.8)*	-276 (11.7)*
High Discretion	-32 (0.9)	-20 (0.6)	157 (4.3)*	-62 (1.9)
Mortality Rate	66 (12.8)*	60 (11.5)*	58 (10.3)*	56 (10.9)*
Mortality Rate*High Dis	8 (1.1)	2 (0.2)	-30 (3.8)*	7 (0.9)
Percent Risk HMO	98 (4.5)*	117 (5.4)*	108 (4.6)*	107 (4.9)*
Percent Risk HMO*High Dis	-96 (3.1)*	-154 (4.9)*	-193 (5.8)*	-110 (3.6)*
Percent Buy-in	264 (7.2)*	331 (8.9)*	369 (9.2)*	186 (5.1)*
Percent Buy-in*High Dis	47 (0.9)	1 (0.0)	-280 (4.9)*	112 (2.2)*
Physicians /Capita	87 (4.9)*	19 (1.0)	47 (2.4)*	-60 (4.2)*
Physicians/Capita*High Ds	-13 (0.5)	106 (4.1)*	141 (5.1)*	106 (4.2)*
Adjusted R^2	0.31	0.30	0.22	0.25
Hospital Bed Supply Model				
Constant	-342 (14.0)*	-325 (13.3)*	-334 (12.4)*	-299 (12.4)*
High Discretion	-60 (1.7)	-44 (1.3)	175 (4.6)	-63 (1.8)
Mortality Rate	57 (11.2)*	60 (11.5)*	57 (10.1)*	57 (11.2)*
Mortality Rate* High Dis	5 (0.7)	-0 (0.1)	-33 (4.0)*	5 (0.7)
Percent Risk HMO	112 (5.2)*	117 (5.3)*	111 (4.6)*	112 (5.2)*
Percent Risk HMO*High Dis	-107 (3.5)*	-144 (4.6)*	-195 (5.7)*	-107 (3.5)*
Percent Buy-in	177 (4.9)*	334 (9.0)*	383 (9.4)*	177 (4.9)*
Percent Buy-in*High Dis	140 (2.7)*	36 (0.7)	-248 (4.3)*	140 (2.7)*
Hospital Beds/Capita	7 (0.4)	4 (0.2)	49 (2.2)*	7 (0.4)
Hospital Beds/Capita*High	90 (3.2)*	151 (5.2)*	73 (2.3)*	90 (3.2)*
Adjusted R^2	0.25	0.30	0.19	0.25

* $P < 0.05$

the estimated coefficients for the interaction variable PERCENT BUY-IN* HIGH DISCRETION were statistically significant, but there was no consistency in the signs of the estimated coefficients across models and states.

Since the main focus of the modeling efforts of this chapter are to test hypotheses concerning the sensitivity of discretionary hospital use rates to market supply factors and HMO practice styles, the empirical findings for the estimated coefficients relevant for testing hypotheses Ha and Hb stated earlier in this chapter are summarized in Table 5.6.

The top portion of Table 5.6 pertains to hypothesis Ha regarding the greater sensitivity of high discretion hospital discharge rates to physician and hospital bed supply factors. The relevant estimated regression coefficients are the interaction terms between the HIGH DISCRETION dummy variable and the respective physician or hospital bed supply variable. A positive (negative) estimated coefficient is supportive (not supportive) of the hypothesis Ha. About three-fourths of the models estimated with either the physician or hospital supply variable specified had positive coefficients supportive of Ha, but only about half of these estimated positive coefficients were statistically significant at conventional levels. Only two models had significant negative coefficients estimated for the physician supply/ high discretion interaction term that contradicted the hypothesis. It is also notable that for the models estimated on the larger pooled sample of geographic units from all study states shown earlier in Table 5.5, eight of nine supply factor interaction variables had positive and significant coefficient estimates.

Overall, the empirical findings are moderately supportive of Ha and the hypothesized greater sensitivity of high discretion hospital discharge rates to physician and bed supply factors. Comparisons among the alternative discretion classifications in Table 5.6 suggest that the

Table 5.6: Counts of Models with Estimated Coefficients Supportive or Not Supportive of Study Hypothesis Ha and Hb and their Significance Levels

			Discretion Classification System				
			DCG Model	Anderson Patient Variation	Anderson Physician Discretion	Roos Index of Discretion	Total
Ha Physician Supply	Ha Supported	Sig.	0	1	3	3	7
		Not Sig.	2	4	0	1	7
	Ha Not Supported	Sig	1	0	1	0	2
		Not Sig.	2	0	1	1	4
Ha Hospital Beds Supply	Ha Supported	Sig.	1	[illegible]	[illegible]	2	8
		Not Sig.	2	1	2	[illegible]	7
	Ha Not Supported	Sig.	0	0	0	0	0
		Not Sig.	1	0	2	1	5
Hb HMO Practice Styles	Hb Supported	Sig.	4	[illegible]	[illegible]	6	17
		Not Sig.	[illegible]	[illegible]	0	[illegible]	15
	Hb Not Supported	Sig.	0	0	0	0	0
		Not Sig.	4	0	0	4	8

Anderson patient variation and Roos index of discretion classifications yielded more coefficient estimates that were supportive of the hypothesis Ha than the other two systems. However, the results do not suggest the relative superiority of any one discretion classification system over all others with respect to supportive empirical results.

The bottom portion of Table 5.6 contains a summary of estimated coefficients for the interaction variable PERCENT RISK HMO* HIGH DISCRETION. Under Hb, a negative estimated coefficient is expected under the premise that HMOs are more successful in controlling discretionary hospital use. Examination of the summarized findings suggest stronger support for the study hypothesis Hb than was found for hypothesis Ha. Eighty percent of the 40 models estimated had estimated coefficients that were consistent with expectations under Hb. A little more than half of these coefficients with expected signs were statistically significant, and no estimated coefficients were both significant and of incorrect sign under Hb. Furthermore, the estimated coefficients were of correct sign and significant in all eight models estimated with the larger pooled sample of geographic units.

Overall, the empirical results are supportive of the hypothesis Hb. Discretionary hospital discharge rates are suggested to be lower in geographic areas where a greater fraction of beneficiaries are enrolled in Medicare risk HMOs. However comparisons of the empirical findings among the alternative discretion classification systems do not suggest that any particular discretion classification system stands out from the others with respect to empirical support of Hb. On one hand, both of Anderson's classification systems produced coefficient estimates that were uniformly of expected signs under hypothesis Hb. On the other hand, the Roos index of discretion had more models with significant estimates of expected signs than the Anderson classifications. But there were four models where the coefficients were not significant and the signs were incorrect.

5.4 DISCUSSION

In this chapter, we sought to test the general validity of the alternative discretion classification systems through empirical tests of two deductive hypotheses about expected differences in the sensitivity of discretionary and nondiscretionary hospital discharge rates to local differences in the supply of physicians and hospital beds, and to local differences in the share of beneficiary populations enrolled in Medicare risk HMOs. It was hypothesized that discharge rates for high discretion hospitalizations would be more sensitive to both physician and hospital bed supply, and to risk HMO market penetration level, than low discretion hospital discharge rates. Simple econometric models were estimated employing geographic areas as units of observation for each state and discretion classification system to test the hypotheses advanced.

The estimated parameters of the regression models lent moderate or fairly strong empirical support for each of the study hypotheses. Health status-adjusted discharge rates for hospitalizations classified as high discretion were generally shown to be more highly correlated with levels of physician and hospital bed supply than low discretion hospitalizations for all four discretion classifications tested. Furthermore, for all discretion classifications, adjusted hospital discharge rates for high discretion hospitalizations were generally shown to be more inversely correlated with Medicare risk HMO market penetration levels than low discretion hospitalizations.

Small area variations research has focused much attention on empirically identifying hospitalizations and surgical procedures where there are high levels of geographic variability, presumed to be associated with physician uncertainty about outcomes of alternative courses of treatment. It has not devoted much, if any, attention to the subsequent question of whether hospitalizations which are distinguished because of their high or low level of geographic variability, are determined by different demand and supply factors than other hospitalizations. The analyses of this chapter are the first to date to estimate multivariate models of factors associated with geographic variations in hospital use rates for

subgroups of hospitalizations differentiated by discretion level, and to use the empirical findings to test deductive hypotheses associated with the concept of discretionary hospitalizations itself.

The analyses of this chapter provide some important empirical evidence in support of the validity of the general concept of discretion levels of hospitalizations. Unfortunately, the empirical analyses did not discriminate very well among alternatives discretion classifications to lead one to favor any particular classification system over all others. In contrast to the correlation analyses reported in Chapter 4, the empirical performance of the Roos index of discretion did not stand apart from the others. However, the generally weaker and more inconsistent empirical findings for the DCG discretion classification system are consistent with findings from other chapters.

SUMMARY AND POLICY IMPLICATIONS

Chapter 6

6.1 OVERVIEW

The major goal of this study was to investigate the validity of the discretionary component of the DCG risk classification system through a comprehensive empirical analysis of differences in the diagnostic composition of hospitalizations of Medicare beneficiaries who receive care in settings where medical practice styles are likely to differ. Employing the DCG discretion classifications and the discretion classification systems of Anderson, et al. (1989) and Roos, et al. (1988) as alternative means of distinguishing among high, low, and moderate discretion hospitalizations, the study goals were addressed by subjecting two main conceptual hypotheses to empirical testing:

- **"high discretion" hospital admissions will account for a significant part of the higher overall hospital admission rates of FFS Medicare beneficiaries relative to risk Medicare HMO enrollees in the same geographic markets;**

- **"high discretion" hospital admissions will account for a significantly greater part of overall hospital admission rates in high-use geographic markets than in low-use geographic markets.**

The comparative empirical performance of alternative discretion classifications provides objective information for evaluating the validity of specific DCG discretion classifications as well as the merits of using the general concept of nondiscretionary hospitalizations to ensure fairness of prior use risk classifications among settings where medical practice styles vary.

6.2 PRINCIPAL STUDY FINDINGS

The principal study findings can be summarized in several areas related to the main study hypotheses. Each of these areas are discussed in turn. The broader implications of these findings are discussed afterward.

HMO-FFS Differences in Discretionary Hospitalizations

Descriptive analyses of the diagnostic composition of over 2.6 million Medicare HMO and FFS hospitalizations for four states in Chapter 2 provided little empirical support for the main study hypothesis, namely, that a significant portion of lower Medicare risk HMO hospital use rates are associated with their success in reducing discretionary hospitalizations as defined by the DCG discretion classes. Rather, the diagnostic composition of Medicare risk HMO and FFS hospitalizations appears to be relatively invariant with respect to both DCG risk class and discretion score classification. As similar results were found using the alternative discretion classifications of Anderson, et al. (1989) and Roos, et al. (1988) in Chapter 3, the DCG model findings are unlikely to be the result of the misclassification of a few diagnoses among discretion classes. These empirical findings are consistent with Luft's (1978) early findings of across the board lower commercial HMO use rates for both discretionary and nondiscretionary hospitalizations. While the basis for this apparent relative invariance in the discretionary composition of Medicare risk HMO and FFS hospitalizations is not well understood, it has implications toward the treatment of discretionary hospitalizations in prior use risk adjusters that will be discussed later in this chapter.

Geographic Variations in Hospital Use and Discretionary Hospitalizations

Geographic analyses were conducted to test the main study hypotheses regarding the relative geographic variability of high versus low discretion hospitalizations, and the degree to which higher overall rates of hospital use are attributable to excessive high discretion hospitalizations. Correlation and regression analysis findings for the DCG and alternative discretion classifications reported in Chapter 4 suggest higher use rate geographic areas tend to have a greater share of hospitalizations classified as high discretion, but higher overall hospital use rates are not largely attributable to excessive discretionary hospital use.

In fact our empirical analyses suggest that higher overall hospital use rates among geographic areas are more strongly associated with a residual group of hospitalizations for all conditions not classified as "low discretion." Geographic areas with high or low overall hospital use rates appear to have relatively similar absolute hospital use rates for low discretion hospitalizations. A low fraction of "low discretion" hospitalizations in a geographic areas appears to serve as a much more effective marker for distinguishing high overall use rate geographic areas than a high fraction of "high discretion" hospitalizations.

Comparisons of the relative empirical performance of the DCG and Anderson, et al. (1989) discretion classifications in the correlation and regression analyses of Chapter 4 indicate only very modest differences among alternative discretion classifications derived from physician a priori ratings of principle diagnoses of hospitalizations. The empirically-derived Roos, et al. (1988) index of discretion stood apart from the DCG and Anderson discretion classifications in terms of superior empirical performance. Nevertheless, the empirical findings from the Roos index of discretion could hardly lead one to conclude that higher overall hospital use rates were predominantly the result of highly discretionary hospital use. The relative geographic invariance of the composition of hospitalizations under the Roos index of classification suggests that there are significant limitations associated with distinguishing the discretion level of hospitalizations from the principal discharge diagnostic information.

Finally the multivariate regression analyses in Chapter 5 provide important empirical evidence that all of the discretionary classifications tested were "on the right track" in distinguishing among high and low discretion hospitalizations. Hospital discharge rates for high discretion hospitalizations were shown to be more sensitive to levels of physician and hospital bed supply than for low discretion hospitalizations. Furthermore, hospital discharge rates for high discretion hospitalizations were found to be significantly lower in geographic areas with greater concentrations of Medicare risk HMO

enrollment. Nevertheless, the differences found were fairly modest as most of the variance in hospital discharge rates could not be explained by the discretion level of hospitalizations.

6.3 POLICY IMPLICATIONS FOR PRIOR USE MODEL RISK CLASSIFICATION

The study findings have direct implications toward the merits of approach Ellis and Ash (1995) took for reducing the potential biases of medical practice style differences on prior use risk classifications. Their approach of excluding highly discretionary prior hospitalizations for assigning Medicare beneficiaries to higher risk (and higher payment rate) risk cells can only be effective if the hospitalizations singled out as being highly discretionary can be validly distinguished from other less discretionary hospitalizations. If "nondiscretionary hospitalizations" are empirically defined as those for which health status-adjusted admission rates are relatively invariant between HMO and FFS populations or among geographic areas, our empirical analyses indicate that it is very difficult to distinguish many such hospitalizations with only information on principal diagnosis. It is equally difficult to use only primary diagnoses to distinguish high discretion hospitalizations which account for a significant portion of the higher hospital admission rates observed for some geographic areas and for Medicare FFS beneficiaries relative to HMO enrollees.

In general these study findings suggest that the discretionary component of the DCG model of Ellis and Ash (1995) is unlikely to serve its original intended purpose very well. It appears that either systematic HMO-FFS practice style differences have little impact on diagnostic composition of hospitalizations, or the measurements of discretion in the DCG model of Ellis and Ash (1995), Anderson, et al. (1989), or Roos, et al. (1988) do not have sufficient validity to warrant their use for excluding certain "high discretion" hospitalizations for higher risk classification. Our findings provide empirical support for Ellis, et al.'s (1995) decision to abandon the formal concept of physician discretion from forerunner DCG models in the development of HCC models. Our findings do not

impart any implications toward the merits of expanding the sources of diagnostic information for HCC risk classification to include ambulatory claims, however. Potential concerns over biases associated with HMO-FFS practice style differences or provider gaming behavior with HCC models should be focused on diagnostic assignment derived from outpatient and/or physician utilization claims.

6.4 IMPLICATIONS FOR SMALL AREA ANALYSIS OF GEOGRAPHIC VARIATIONS

The study findings have some important broader implications for studies of geographic variations in hospital use as well. While concerns over the reliability of primary payer fields in the hospital discharge data precluded direct analyses of Medicare risk HMO and FFS differences in hospital utilization rates, our limited use of Medicare risk HMO enrollment data in the empirical geographic analyses of hospital use rates for combined Medicare risk HMO and FFS beneficiary populations suggests that estimation of disaggregated HMO and FFS hospital use rate models would be likely to produce some valuable empirical insight about the impacts of population health status differences on variations in hospital utilization rates. Our descriptive analyses indicated that risk HMO-FFS differences in estimated age-sex adjusted mortality rates were much greater than geographic differences in mortality rates for the combined HMO and FFS study populations, particularly in California and New York. Given the strong correlations we found between hospital use rates and the mortality rates for the combined HMO and FFS populations of geographic areas, the greater dispersion in mortality rates among Medicare risk HMO versus FFS beneficiaries may provide a means for distinguishing population health status effects on geographic variations in hospital use rates that are muted by the aggregation of HMO and FFS data.

A pilot study employing hospital discharge data in which Medicare risk HMO hospital discharges are accurately distinguished by matching Medicare enrollment files to state hospital discharge records could be done for a state such as California or Florida with larger Medicare risk

HMO enrollments. If disaggregated geographic HMO-FFS models of hospital use rates were found to impart added insight about population health status impacts on hospital use rates, such models may provide a potential empirical tool for monitoring problems of access and/or enrollment selection bias in Medicare risk HMOs. Medicare risk HMOs with much lower hospital use rates than predicted under such geographic models might be flagged for further scrutiny to assess problems of access and/or enrollment selection bias.

REFERENCES

Adamache, K., and L. Rossiter, "The Entry of HMOs into Medicare: Implications for TEFRA's Mandate," Inquiry 23, 1986, 349-364.

Ash, A., Porell, F., Gruenberg, L., Sawitz, E., and A. Beiser, "Adjusting Medicare Capitation Rates Using Prior Hospitalization," Health Care Financing Review 10:4, 1989, 84-95.

Ash, A., Ellis, R. and L. Iezzoni, "Clinical Refinements to the Diagnostic Cost Group Model," Final Report to the Health Care Financing Administration, Health Care Research Unit, Boston University School of Medicine, 1989.

Blumberg, M.S., "Regional Differences in Hospital Use Standardized by Reported Morbidity," Medical Care 20: (9) 1982 931-944.

Blumberg, M.S., "U.S. Inter-Area Variation in Age-Adjusted Health Status," Unpublished Working Paper, 1985.

Brown, R. "Biased selection in the Medicare competition demonstrations," Final Report to the Health Care Financing Administration, (Princeton, NJ: Mathematica Policy Research) 1988.

Brown, R., Bergeron, J., Clement, D., Hill, J., and S. Retchin, "The Medicare Risk Program for HMOs: Final Summary Report on Findings from the Evaluation," Final Report to the Health Care Financing Administration, (Princeton, NJ: Mathematica Policy Research), 1993.

Chassin, M.R., "Explaining Geographic Variations: The Enthusiasm Hypothesis," Medical Care 31:5, 1993, YS37-44.

Chassin, M.R., Kosecoff, J., Park, R.E., Winslow, C., Kahn, K., Merrick, N., Keesey, J., Fink, A., Solomon, D., and R. Brook, " Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services?" Journal of the American Medical Association 258:18, 1987, 2533-2537.

Deacon, R., Lubitz, J., Gornick, M., and M. Newton, "Analysis of Variations in Hospital Use by Medicare Patients in PSRO Areas: 1974-1977,"Health Care Financing Review 1:1 1979, 79-107.

Diehr, P., K. Cain,, Z. Ye, and F. Abdul-Salam, "Small Area Variation Analysis: Methods for Comparing Several Diagnosis-Related Groups," Medical Care 31:5, 1993, YS45-53.

Eggers, P.W., Prihoda, R. "Pre-enrollment reimbursement patterns of Medicare beneficiaries enrolled in at-risk HMOs," Health Care Financing Review, 4, 1982 55-73.

Ellis, R., and A. Ash, " Refinements to the Diagnostic Cost Group Model," Inquiry 32, Winter 1995, 1-12.

Ellis, R., Pope, G., Iezzoni, L., Ayanian, J. , Bates, D., Burstin, H., and A. Ash, "Diagnosis-Based Risk Adjustments for Medicare Capitation Payments," Health Care Financing Review 17, (3), Spring 1996, 101-128.

Epstein, A., and E. Cumella, "Capitation Payments: Using Predictors of Medical Utilization to Adjust Rates," Health Care Financing Review 10, 1988.

Folland, S. and M. Stanno, "Small Area Variations: A Critical Review of Propositions, Methods, and Evidence," Medical Care Review 47:4, 1990, 419-465.

Gornick, M., "Trends and Regional Variations in Hospital Use under Medicare," in D.L. Rothberg (ed.), Regional Variations in Hospital Use, (D.C. Heath and Company: Massachusetts), 1982, 131-184.

Griffith, J., Wilson, P., Wolfe, R., and D. Bischak, "Clinical Profiles of Hospital Discharge Rates in Local Communities," Health Services Research 20:2 1985, 131-151.

Gruenberg, L., Wallack, S., and C. Tompkins, "Pricing Strategies for Capitated Delivery Systems," Health Care Financing Review 7: 1986, 35-44.

Gruenberg, L., Tompkins, C., and F. Porell, "The Health Status and Utilization Patterns of the Elderly: Implications for Setting Medicare Payments to HMOs" in R. Scheffler and L. Rossiter (eds.), Advances in Health Economics and Health Services Research 10, 1989, 41-74.

Health Care Financing Administration, "Medicare and Medicaid Statistical Supplement," Health Care Financing Review 1995.

Hornbrook, M., Bennett, M., and M. Greenlick, "Diagnostic Cost Groups: Validation from an HMO's Experience," Paper Presentation, American Public Health Association, 1988.

Kunkel, S., and C. Powell, "The Adjusted Average Per Capita Costs Under Risk Contracts with Providers of Health Care," Transactions of the Society of Actuaries 18, 1981, 57-66.

Leape, L., Park, R., Solomon, H., Chassin, M., Kosecoff, J., and R. Brook, "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services? A Study of Three Procedures," Journal of the American Medical Association 258:18 November 1987, 2533-37.

Lubitz, J., Beebe, J., and G. Riley, "Improving the Medicare HMO Payment Formula to Deal With Biased Selection," in R. Scheffler and L. Rossiter (eds.), Advances in Health Economics and Health Services Research 6, 1985, 101-122.

McClure, W. and D. Shaller, "Variations in Medicare Expenditures," Health Affairs 3:2 1984, 120-129.

Luft, H.S., "How Do Health Maintenance Organizations Achieve their Savings? Rhetoric and Evidence," New England Journal of Medicine June 1978, 1336.

McMillan, A., "Trends in Medicare Enrollment in Health Maintenance Organizations: 1986 through 1993," Unpublished paper, 1993.

McPherson, K., Strong, P., Epstein, A., and L. Jones, "Regional Variations in the Use of Surgical Procedures: Within and Between England and Wales, Canada, and the United States of America," Social Science and Medicine 15:3 1981, 273-288.

McPherson, K., Wennberg, J., Hovind, O., et al., " Regional Variations in the Use of Common Surgical Procedures: An International Comparison of New England, England, and Norway," New England Journal of Medicine 307:21, 1982, 1310-1314.

National Center for Health Statistics, Monthly Vital Statistics Report 42, (2), August 31, 1993.

Newhouse, J., "Rate Adjusters for Medicare Under Capitation," Health Care Financing Review, Annual Supplement, 1986, 45-55.

Park, R., "Does Inappropriate Use Explain Small-Area Variations in the Use of Health Services," A Reply," Health Services Research 28:4 ,1993, 401.

Paul-Shaheen, P., Clark, J.D., and D. Williams, " Small Area Analysis: A Review and Analysis of the North American Literature," Journal of Health Politics, Policy, and Law 12:4 1987, 741-808.

Porell, F., Welch, W., Tompkins, C., and W. Turner ,"An Empirical Analysis of Alternative Geographic Configurations for Basing Medicare Payments to HMOs" Final Report to HCFA, May 1988.

Porell, F. and W. Turner, "Biased selection under an experimental enrollment and marketing Medicare HMO broker," Medical Care, 28(7), 1990a. 606-615.

Porell, F. and W. Turner, "Biased selection under the Senior Health Plan prior use capitation formula," Inquiry 27: 1990b, 39-50.

Porell, F., Tompkins, C., and W. Turner, " Alternative Geographic Configurations for Medicare Payments to HMOs," Health Care Financing Review 11:3, 1990, 17-30.

Porell, F. and C. Tompkins, "Regional Variations in Rates of Medicare Hospital Admissions and the Diagnostic Composition of Hospital Episodes," Unpublished Working Paper, 1993.

Porell, F. and S. Wallack, "HMO Medicare Risk Contracting: The Determinants of Market Entry," Health Care Financing Review 12 (2), 1990, 75-85.

Porell, F., Tompkins, C., Perales, P., and L. Gruenberg, "Employer Insured Beneficiaries: Implications for HMOs and Medicare," Final Report to the Health Care Financing Administration, Institute for Health Policy, Brandeis University, May 1991.

Riley, G., Rabey, E., and J. Kasper, "Biased Selection and Regression to the Mean in Three Medicare HMO Demonstrations: A Survival Analysis of Enrollees and Disenrollees," Medical Care 27:4 1989.

Riley, G., Lubitz, J., and E. Rabey, "Enrollee Health Status Under Medicare Risk Contracts: An Analysis of Mortality Rates," HCFA Report, July, 1989.

Roos, N.P., Wennberg, J., and K. McPherson, "Using Diagnosis-Related Groups for Studying Variations in Hospital Admissions," Health Care Financing Review 9:4 1988, 53-62.

Rossiter, L., and K. Adamache, "The Determinants of Geographic Variation in Medicare Costs: Market Forces in a Broader Context," Final Report to the Health Care Financing Administration, Williamson Institute for Health Studies, Medical College of Virginia, 1988.

Shwartz, M., Ash, A., Anderson, J., Iezzoni, L., Payne, S. and Restuccia, J., "Small Area Variations in Hospitalization Rates: How Much You See Depends on Ho You Look," Medical Care, Vol. 32, No. 3, pp. 189-201.

U.S. General Accounting Office, "Medicare issues raised by Florida health maintenance organizations," Report to Congress. GAO/HRD-86-97. Washington, DC: U.S. Government Printing Office, 1986.

Ward, J., "Hierarchical Grouping to Optimize an Objective Function," American Statistical Association Journal 58: 1963.

Wennberg, J., "Dealing With Medical Practice Variations: A Proposal for Action," Health Affairs 1985; 3:6.

Wennberg, J., and A. Gittelsohn, "Variations in Medical Care Among Small Areas," Scientific American 246:4 1982, 120-134.

Wennberg, J., Barnes, B., and M. Zubkoff, "Professional Uncertainty and the Problem of Supplier Induced Demand," Social Science and Medicine 16:7, 1982, 811-824.

Wennberg, J., Freeman, J., and W. Culp, "Are Hospital Services Rationed in New Haven or Over-utilized in Boston?" Lancet 1 1987, 1185-1188.

Wennberg, J., Freeman, J., Shelton, R., and T. Bubolz, "Hospital Use and Mortality Among Medicare Beneficiaries in Boston and New Haven," The New England Journal of Medicine 1989; 321:1168.

Yip, W., and H. Luft, "Border Crossing for Hospital Care and its Implications for the Use of Statewide Data," Social Science and Medicine 36:11, 1993, 1455-1465.

APPENDIX A

METHODOLOGY FOR DEFINING GEOGRAPHIC UNITS

A.1 INTRODUCTION

The purpose of this appendix is to provide a more detailed discussion of the conceptual basis and empirical construction of the geographic units employed in the study. The following section contains a brief discussion of the conceptual rationale for defining geographic units on the basis of patient patterns of hospital choice, and how this approach differs from more conventional approaches to defining hospital service areas. The third section illustrates the mechanics of the clustering algorithm for defining the study geographic units through a simple hypothetical example with patient origin data. The fourth section contains a detailed discussion of the methodology employed in the construction of hospital choice area geographic units in this study.

A.2. BACKGROUND

The delineation of geographic units is obviously an important element in the analysis of geographic variations in medical care utilization. Ideally geographic units should be large enough to contain sufficient population for the reliable measurement of health service utilization, but they also should be small enough such that the utilization of all residents reflect a common style of medical practice. Most of the hospital service area methodologies in the literature (e.g., Garnick et. al, 1987) essentially group together the service areas of hospitals that draw patients from similar zip code areas through a chaining process. Hospital A may be grouped with hospital B if they both draw a significant number of patients from some subset of zip codes. Hospital C may then be added to an areawide hospital service area, comprised of hospitals A and B, if it draws a significant number of patients from some of the same zip codes as hospital B does. At the same time, however, there may be no overlap at all in the zip codes where hospital A and C draw patients. Such grouping methodologies essentially produce geographic units which are approximations to a closed system. That is,

they produce geographic units in which patient origin flows associated with hospitals located outside of the broad area are minimal. While this feature is desirable, the chaining process can produce geographic units comprised of sub-areas where patients in one sub-area don't use any of the hospitals used by patients in another sub-area. If the medical practice styles of physician and hospital providers vary spatially, it is more reasonable to posit that patients in sub-areas using common providers should be influenced similarly by medical practice styles and should be clustered together in the same geographic unit.

A study conducted by Porell et. al (1991), invoked the concept of similar hospital choice to cluster sub-areas to form larger geographic units which would best reflect the underlying geographic variations among smaller sub-areas. The underlying premise for the clustering algorithm is that zip codes where patients tend to use the same hospitals are more likely to exhibit utilization patterns influenced by common medical practice styles. Under this premise, the clustering of zip codes with very similar patterns of usage among hospitals should produce larger geographic units that are relatively homogenous in terms of within-unit variance of zip code level utilization or expenditure rates. Porell et. al (1991) found hospital choice areas produced by this algorithm to have greater internal homogeneity (i.e., smaller, relative, within-unit variance of Medicare reimbursements per capita) than five alternative geographic units comprised of multiple 5-digit zip codes.

A.3 Hospital Choice Clustering Algorithm

The "hospital choice area" clustering algorithm of Porell et. al (1991) is similar to Ward's hierarchical clustering algorithm (Ward, 1963) which has been employed by Schwartz et. al (1994) for defining market areas for analysis of geographic variations. Both algorithms measure hospital use patterns in terms of the proportions of patients discharged from various hospitals. The main difference between the algorithms lies in the measurement of similarity in these hospital use patterns. In Ward's (1963) algorithm greater similarity in hospital use for a pair of zip codes is measured in terms of smaller sum of squared differences between the zip code proportions of patients discharged to each hospital. In the Porell et. al (1991) algorithm, greater

similarity is measured in terms of the overlap in the proportions of patients discharged to various hospitals. As discussed later, overlap is measured directly as the proportion of the populations of a pair of zip codes using the same hospitals, where zip code population sizes are equal. This measure is similar but not equivalent to absolute deviations of differences in proportions. Otherwise, the two algorithms operate similarly. In the first step of both algorithms, the two zip codes with the most similar pattern of hospital choices are grouped together regardless of whether or not they are spatially contiguous. Similarity measures among all pairs of areas are then recomputed with respect to the newly aggregated area and the next pair of units with the greatest similarity are grouped in the second step. This process continues with a pair of units grouped at each subsequent step of the algorithm until a stopping criterion is met, or until all zip codes are aggregated into a single large area.

The logic of the clustering algorithm can be more easily understood through a hypothetical example of its application. The objective of the methodology is to group together areas where the revealed hospital choices of residents are most similar. Hospital patient origin data are used to approximate the broader concept of similar provider choice. Specifically, we wish to group together those areas where residents of both areas use the same hospitals. The following hypothetical example of five areas and four hospitals should help in illustrating the basic logic in the grouping algorithm.

Consider the patient origin table of hospital discharges from five residence areas (1,2,3,4,5) to four hospitals (A,B,C,D) in Table A.1. To normalize for differing population size, we compute the shares or fractions of individuals going to different hospitals from each residence area as in Table A.2. The normalized patient origin matrix shows the patient origin choice patterns for a set of areas that are of similar size. What we now wish to measure is the extent to which there is overlap in hospital choices of the residents in any two pairs of areas. Overlap can be simply defined as the fraction of people (controlling for population size) who go to the same hospitals from any pair of residence areas. First consider residence areas one and two. Half of residence area one people go to hospital A. Only ten percent of residence area two people go to that same hospital. If both residence areas had 100 hospital admissions (arising from 100 different persons), ten persons

Table A.1: Hypothetical Patient Origin Table with Five Residence Areas and Four Hospitals

Area/Hospital	A	B	C	D	TOTAL
1	500	250	200	50	1000
2	40	200	80	80	400
3	90	60	90	60	300
4	180	60	0	360	600
5	140	30	30	30	200

Table A.2: Normalized Patient Origin Matrix of the Proportions of Residence Area Hospital Discharges to Each Hospital

Area/Hospital	A	B	C	D	TOTAL
1	0.50	0.25	0.20	0.05	1.00
2	0.10	0.50	0.20	0.20	1.00
3	0.30	0.20	0.30	0.20	1.00
4	0.30	0.10	0.00	0.60	1.00
5	0.70	0.10	0.10	0.10	1.00

in residence area one would have chosen the same hospital A as ten persons who chose hospital A from residence area two. The remaining 40 individuals in residence area one who chose hospital A had no counterpart in residence area two choosing hospital A. The overlap in the hospital admission pattern can be seen as the minimum of the shares of hospital admissions to any particular hospital. Thus, there is an overlap of 10 hospital admissions when admissions to hospital A are considered.

Half of residence area two hospital admissions go to hospital B. Only 25 percent of residence area one admissions go to hospital B. The overlap in this choice pattern is 25 percent, or the minimum of 50 percent and 25 percent. Twenty-five persons from each 100 patient residence areas were admitted to the same hospital, B.

For hospital C, the hospital choice overlap is 20 percent since both patient admission shares are 20 percent. Twenty people from both residence areas each went to hospital C. For hospital D, the overlap is five percent, or the minimum of five percent and 20 percent. Five persons from residence area one went to the same hospital D as five other persons from residence area two.

Summing the overlap over the four hospitals A through D, the total overlap in the pattern of hospital choice for residence areas one and two is: $(10\% + 25\% + 20\% + 5\%) = 60\%$. What this means is that 60 percent of the hospital choice patterns overlap for residence areas one and two. If there were truly 100 hospital admissions from each residence area, 60 persons from each residence area had a counterpart in the other residence area going to the same hospital.

Before computing overlap measures for all pairs of residence areas, it may be important to discuss the two extremes of the overlap measure. First, if the fractions of people choosing different hospitals for any pair of residence areas are identical, the overlap will be 100 percent. Every person from one residence area has a counterpart in another residence area choosing the same hospital. An overlap of 100 percent does not require that all persons in two areas be admitted to a single hospital. This can arise if all persons are admitted to a single hospital, or when admissions are to a large number of different hospitals. In the latter case, the proportion of individuals going to each hospital would have to be identical for complete choice overlap.

Alternatively, if all individuals from one residence area are hospitalized in a different hospital or different set of hospitals than those in a second residence area, the overlap measure will be zero.

The more similar are hospital choice patterns, the greater will be the choice overlap measure (with an upper bound of 100%) when computed in this way. This overlap measure is superior to conventional distance measures employed in cluster analysis for the purposes at hand. A more conventional distance measure, such as the mean deviation or squared deviation between the proportions of individuals admitted to various hospitals, could lead to erroneous conclusions about similarities in actual patient hospital use. This is because "similar hospital use" under a conventional distance measure exists not only when patients from two areas use the same hospitals, but also when patients do not use the same hospitals. Suppose that there were ten hospitals in a state. If 100 percent of residence area one persons were admitted to hospital A, and 100 percent of residence area two persons were admitted to hospital B, the proposed overlap measure developed here would be zero. A simple mean deviation distance measure for these two areas would be ten percent, however. There would be no deviation between the two residence areas in the fractions of patients admitted to the remaining eight hospitals in the state because no patients from either residence area were admitted to the remaining eight hospitals.

Table A.3 contains the overlap measures for the matrix of unique pairs of residence area combinations for the five residence areas and four hospitals. Comparisons of units with themselves are always 100 percent by definition. A hierarchical aggregation approach was used to cluster areas into choice areas with these overlap measures. What this means is that choice areas are formed incrementally by grouping together pairs of residence areas with maximum overlap in hospital choice distributional patterns. Once a pair of residence areas are clustered together from N initial areas, the overlap matrix is recomputed for the remaining N-1 areas (i.e., the remaining N-2 original areas plus the first paired cluster). Then a pair of these N-1 areas are chosen for clustering on the basis of maximum overlap.

Under this hierarchical aggregation approach, either Areas one and three, or Areas one and five could

Table A.3: Pairwise Overlap in Distribution of Hospital Discharges Among Residence Areas

Areas	1	2	3	4	5
1	100%	60%	75%	45%	75%
2		100%	70%	40%	40%
3			100%	60%	60%
4				100%	50%
5					100%

be grouped together first, since each has the maximum overlap of 75 percent. Grouping together areas one and five first, a new residence area 1-5 is defined to be now comprised of areas one and five together. Patient origin data would be aggregated together for areas 1 and 5, and the normalized patient share data would be now computed for the new aggregated area 1-5.

The choice pattern overlaps would now be recomputed for the four remaining areas (area 1-5, two, three, and four). From the new overlap matrix, areas two and three would be aggregated together to form area 2-3. This now leaves three payment areas (1-5,2-3, and four). The hierarchical aggregation continues until a stopping criterion is defined and satisfied. Otherwise, all areas will eventually be grouped together into a single payment area under a hierarchical clustering algorithm.

A.4. OPERATIONAL METHODOLOGY

Under the algorithm described above, a hierarchical grouping approach is employed where individual 5-digit zip codes, or previously clustered groups of 5-digit zip codes, are grouped together by choosing the maximum overlap between any two pairs of areas. To implement such an algorithm with actual patient origin data, a variety of operational issues must be considered. Two operational questions of importance pertain to the imposition of the stopping criteria for the algorithm, and the imposition of contiguity constraints. At what point should individual zip codes, or zip code clusters be maintained as separate hospital choice areas versus aggregating them further into a smaller number of geographic units? Setting stopping criteria involves trade-offs between the population size of the geographic units and the number of geographic units serving as units of observation for geographic analysis. Contiguity constraints are important in delineating geographic study units, if for no purpose other than simplicity in geographic configuration. Since there was no contiguity constraint incorporated in the actual clustering algorithm, spatially cohesive geographic units could not be ensured without allowing for adjustments to the geographic configurations produced by the clustering algorithm.

A.4.1 Requirements for Study Spatial Units

The actual process of defining geographic units was developed with the purpose of defining units to best meet some basic requirements of study geographic areas. First, it was important that the individual geographic units be spatially cohesive and all units together be spatially exhaustive. Second, a minimum threshold population of about 4,000 Medicare beneficiaries was sought for each unit since this population size is sufficient for estimating Medicare hospital use rates with standard errors that are less than five percent of the sample mean. Third, given that there are fewer Medicare risk HMO enrollees than FFS beneficiaries generally, to maximize the number of geographic study units with at least 1,000 Medicare risk HMO enrollees. Finally, we sought to generally maximize the number of geographic units to increase the study sample size for geographic analyses. Given the conflicting nature of some of these objectives, the study geographic units were delineated through an interactive process that was strongly guided by the assignments produced by the clustering algorithm.

A.4.2 Geographic Unit Delineation

The interactive process of geographic unit delineation consisted of several steps. In the first step, the clustering algorithm was run for each state, or subregion in the case of New York and California, until there were 100 separate clustered geographic units remaining. In the second step, the algorithm assignments of individual zip codes in each cluster were recorded in the form of tree diagrams showing which zip codes were added to a cluster, and at what iteration of the algorithm these additions occurred. These tree diagrams were used to split the largest, most populated clusters into smaller ones with roughly 1,000 or more discharges by selectively "stopping" the algorithm to preclude the formation of the larger cluster. The third step of the process involved the mapping of the algorithm clusters and manual reallocation of problem zip code assignments. The original algorithm clusters and the split subarea clusters were then mapped to see whether they were spatially cohesive. While the great bulk of the clustering algorithm assignments were spatially

cohesive, a relatively small number of clusters had "holes" comprised of one or more zip codes assigned to some other geographic cluster, or had one or more spatial outlier zip codes, or "islands," that were not contiguous to the main cluster of zip codes. These problem zip codes were then reallocated under a simple set of decision rules to "fix" problems of holes and spatial outliers in the zip code cluster assignments to yield the final set of study geographic units. These steps will be explained more thoroughly below.

Step 1: Application of the hospital choice clustering algorithm

Before the actual application of the clustering algorithm in each of the study states, all 5-digit "point zip codes" were assigned to spatial 5-digit zip codes using data from commercial vendors. Point zip codes are 5-digit zip codes assigned to large facilities (e.g., universities) or residential developments which are spatially concentrated, but are too small to be spatially mapped. Each point zip code is located within a larger spatial zip code. Hospital discharge data with resident zip codes that were point zip codes were recoded as the spatial 5-digit zip code in which the point zip code was located before the clustering algorithm was initiated. Table A.4 shows the number of spatial 5-digit zip codes for each state or region after the assignment of point zip codes to spatial zip code areas. The clustering algorithm was run separately for the states of Massachusetts and Florida, and for north and south regions of California and New York. The clustering algorithms were stopped at the iteration number when there were 100 geographic units left as spatial units. While the bulk of these 100 geographic units were comprised of clusters of 5-digit zip codes, in each state, there were around ten or so individual 5-digit zip codes that remained unmerged with any other zip codes. Most of these unmerged zip codes had very few hospital discharges (i.e., less than 30 discharges). Table A.1 also contains data about the overlap in hospital use distributions at the last iteration of the clustering algorithm. The degree of overlap in hospital use patterns at the final iteration varied among states and subregions. It was highest in the southern portion of New York (61.9%) and in Massachusetts (60.6%), and was lowest in the norther portion of California (38.0%) and Florida (44.1%). Since clusters are formed by merging zip codes with the greatest

Table A.4: Summary of Zip Code Assignments to Geographic Units and Assignment Problems

State/ Region	Spatial 5-digit zip codes at iteration 0	Overlap in hospital discharge distribution at final iteration	Geographic Units after Adjustments to Algorithm Assignments	Percent of 5-digit spatial zip codes with problematic assignments	Percent of hospital discharges in problematic assignments
California (North)	724	38.0% iteration #624	148	3.2%	0.1%
California (South)	562	49.6% iteration #462	148	4.1%	1.6%
Florida	797	44.1% iteration #697	196	0.6%	0.2%
Massachusetts	461	60.6% iteration #361	69	0.6%	0.1%
New York (North)	985	49.5% iteration #885	136	2.1%	1.2%
New York (South)	390	61.9% iteration #290	91	6.6%	1.8%

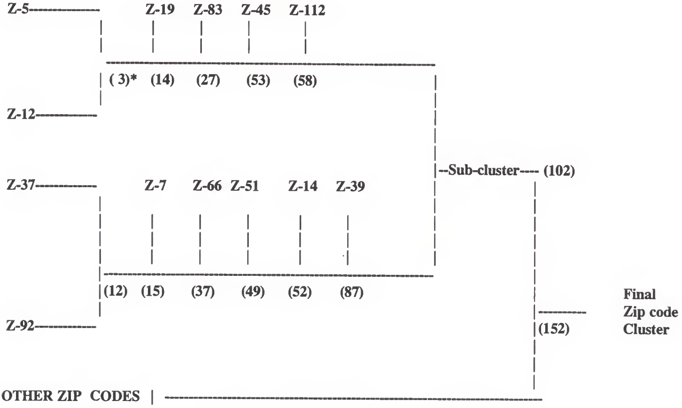
overlap at each iteration of the algorithm, these differences in overlap at the final iteration largely reflect differences in the number of iterations required to reduce the number of geographic units to 100.

Step 2: Decomposition of large zip code clusters

The spatial clusters directly resulting from application of the clustering algorithm varied considerably in terms of their relative population sizes of aged Medicare beneficiaries. Some zip code clusters contained zip codes which together had over 10,000 hospital discharges, while others had less than 1,000 hospital discharges. The splitting of these large zip code clusters into multiple smaller clusters increased the sample size of study geographic units and greatly reduced the maldistribution of their population sizes. Since most of the larger clusters were the result of the merging together of separate large clusters of zip codes at later iterations of the algorithm, the general integrity of the original algorithm assignments could be maintained by precluding the mergers of certain clusters at selected iterations of the algorithm. This approach was chosen over the simpler approach of stopping the clustering algorithm altogether at an earlier iteration, since it allowed the formation of smaller zip code clusters at later iterations of the algorithm resulting in fewer unmerged individual zip codes at the final step.

Potential splits of large zip code clusters were identified by diagraming the algorithm assignments of individual zip codes to clusters. These assignments were recorded in the form of tree diagrams showing which zip codes were added to a cluster, at what iteration of the algorithm these additions occurred, and the number of hospital discharges of the zip codes that were merged together. Figure A.1 illustrates a hypothetical tree diagram of algorithm assignments. Examination of the tree diagram shows that the resulting cluster of 13 zip codes was initiated by the merger of zip codes five and 12 at the third iteration of the algorithm. This was followed by additions of zip codes 19, 83, 45, and 112 to the original cluster at subsequent iterations. A second pair of zip codes 37 and 92 was clustered at the 12th iteration of the algorithm, with the addition of zip codes 7, 66, 51, 14, and 39 at subsequent iterations to form a sub-cluster. The diagram shows that at

Figure A. 1: Hypothetical Tree Diagram of Hospital Choice Area Cluster Assignments



---

* The number in parenthesis is the iteration number of the algorithm where a merger with or addition to cluster takes place.

101 of the algorithm, the two sub-clusters were joined together to form a single larger cluster. At iteration 152, this cluster is merged with a third sub-cluster. The large cluster formed at iteration 152 with 28 zip codes can be split into three smaller clusters of zip codes by selectively precluding the cluster mergers at iterations 101 and 152. By splitting off specific branches of the larger tree diagram at later steps of the algorithm, the general integrity of algorithm assignments is maintained. Each of the three sub-clusters should have greater overlap in their hospital distributions than the single large cluster.

Large zip code clusters were split into smaller clusters only when the smaller clusters had 1,000 or more hospital discharges. Such areas would be expected to have more than 3,000 aged Medicare beneficiaries based on national Medicare hospital discharge rates. Most clusters resulting from these splits exceeded these minimum threshold levels.

Step 3: Adjustments of algorithm cluster assignments

After all algorithm clusters were diagramed and large clusters were decomposed into sub-clusters, each of the resulting clusters were mapped with computer mapping software to assess whether the zip code clusters were spatially cohesive. Over 90 percent of the original algorithm zip code assignments resulted in zip code clusters that were spatially cohesive in all of the study states. The other residual zip code assignments produced two kinds of problems. First, there were "holes" in some geographic unit clusters resulting from zip codes that were never assigned to any cluster at the final iteration of the algorithm or from zip codes assigned to some other cluster. The flip side to the problem of holes in some geographic clusters is the problem of outlier "island" zip codes for other zip code clusters.

The problems of "islands" and "holes" were addressed by manual reassignments of selected zip codes under a set of simple decision rules. In the case of holes, the holes were filled in by reallocating any misallocated zip code to the existing hospital choice area enclosing it. Outliers were dropped from the hospital choice area for which they were an outlier. Hence, under these simple decision rules, an outlier for one

hospital choice area generally served as a hole for another hospital choice area. The exceptions to this general rule involved boundary situations where an outlier was not enclosed by any other hospital choice area. In these situations, the isolated zip code was either retained as a separate geographic unit if it had about 1,000 or more hospital discharges, or allocated to the less populated zip code cluster on either boundary.

Table A.4 contains summary information about the prevalence of problematic algorithm assignments in the study states which required manual reassignment of zip codes and which were not attributable to the decomposition of larger clusters formed by the algorithm. The splitting of large clusters into multiple smaller clusters in some instances created "hole" and "island" problems that would not have arisen had the algorithm assignments been left alone. Since such situations are more the result of a decision to generate a larger sample of comparably-sized geographic units than to faulty clustering algorithm assignments, these are not counted as problematic assignments. Furthermore, unclustered zip codes with sufficient Medicare populations to stand alone as study geographic units were not counted as problematic assignments.

Table A.4 reveals that in the states of Massachusetts and Florida, less than one percent of zip code assignments were problematic as defined above. While still relatively infrequent, problematic zip code assignments were much more prevalent in the states of California and New York. In total, about three percent of algorithm zip code assignments were manually reallocated in these states. The southern portion of New York state had highest prevalence of manual reassignments among all six runs of the clustering algorithm. Since these reallocated zip codes still accounted for less than two percent of hospital discharges in southern New York, it is evident that many were sparsely populated zip codes. While it is possible that the higher rate of problematic zip codes in California and New York was influenced by our splitting of these state into two sub-regions, there was no discernable spatial pattern among reassigned zip codes suggesting an artificial border problem.

A.5 SUMMARY

This appendix contains a discussion of the "hospital choice area" spatial clustering methodology employed for delineation of study geographic units. A mapping of the specific zip codes which comprise each of the study geographic units are not reported here, but are available upon request from the authors.